

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 19:32:18 ; Search time 88.12 Seconds
(without alignments)
5572.190 Million cell updates/sec

Title: US-08-961-083-1

Sequence: 1 TAAATCTAGACATATAA.....ATCCTCAACGACGACAAACA 1999

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	1999	3	US-08-961-083-1
2	1981.4	99.1	2049	3	US-08-481-435-5
3	944.2	47.2	960	2	US-08-245-511-3
4	944.2	47.2	960	2	US-08-600-993A-3
5	117.2	5.9	9100	2	US-08-743-637B-27
6	117.2	5.9	9100	2	US-08-526-840B-27
7	85	4.3	2487	3	US-08-481-435-1
8	62.2	3.1	807	2	US-08-771-716-1
9	62.2	3.1	807	2	US-08-771-716-3
10	62.2	3.1	807	3	US-09-057-720A-1
11	62.2	3.1	807	3	US-09-057-720A-3
12	56.4	2.8	2193	1	US-08-731-716-1
13	56.4	2.8	2193	1	US-08-731-716-3
14	55.2	2.8	428	4	US-09-060-756-206
15	53.8	2.7	1474	4	US-09-508-542-17
16	51.8	2.6	363	4	US-09-060-756-626
17	49.8	2.5	4403765	4	US-09-103-840A-2
18	49.8	2.5	4411529	4	US-09-103-840A-1
19	45.2	2.3	5511	3	US-08-928-361B-2
20	45.2	2.3	7334	3	US-08-928-361B-1
21	41.8	2.1	5163	3	US-08-700-651-1
22	41.8	2.1	5163	3	US-08-928-361B-4
23	41.8	2.1	5318	3	US-08-700-651-2
24	41.8	2.1	5318	3	US-08-928-361B-3
25	40.6	2.0	223	4	US-09-060-756-438
26	36.6	1.8	1172	4	US-08-936-165A-8
27	36.6	1.8	7218	1	US-08-232-463-14

28	36.4	1.8	473	4	US-09-060-756-671	Sequence 671, App
29	35.8	1.8	348	4	US-09-060-756-99	Sequence 99, Appl
30	35.2	1.8	43795	3	US-08-742-185-101	Sequence 101, App
31	34	1.7	2946	4	US-09-175-928-3	Sequence 3, Appl
32	33.8	1.7	1007	4	US-08-858-10668-113	Sequence 113, App
33	33.4	1.7	198	5	PCT-US95-10668-3	Sequence 3, Appl
34	33.4	1.7	198	5	PCT-US95-10668-4	Sequence 4, Appl
35	33.4	1.7	2029	4	US-09-136-574A-46	Sequence 46, Appl
36	33.2	1.7	700	4	US-09-236-097-11	Sequence 11, Appl
37	33	1.7	700	4	US-09-236-097-8	Sequence 8, Appl
38	32.8	1.6	366	4	US-09-060-756-331	Sequence 331, App
39	32.8	1.6	1725	1	US-08-257-073-14	Sequence 82, Appl
40	32.6	1.6	1543	3	US-08-714-918-82	Sequence 82, Appl
41	32.6	1.6	1543	4	US-09-265-315-82	Sequence 82, Appl
42	32.6	1.6	1543	4	US-09-265-315-82	Sequence 82, Appl
43	32.6	1.6	1543	4	US-09-266-417-82	Sequence 82, Appl
44	32.6	1.6	3343	4	US-08-965-762-21	Sequence 21, Appl
45	32.4	1.6	3292	1	US-07-814-964-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-08-961-083-1
Sequence 1, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Chol et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1999 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-1

Query Match 100.0%; Score 1999; DB 3; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAATCTAGACATATAAATCACTGCTGAGTGGTTCGACGCGCGCA 60
Db 1 TAAATCTAGACATATAAATCACTGCTGAGTGGTTCGACGCGCGTCA 60
```

61 TGGCCAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGTTTCTATCGAAGACCA 120
61 TGGCCAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGTTTCTATCGAAGACCA 120
121 TGGCTTCTTCGACACAGGGGGGTTGATACATCCGATCTCTGGAGCTTTCTTGGCCAA 180
121 TGGCTTCTTCGACACAGGGGGGTTGATACATCCGATCTCTGGAGCTTTCTTGGCCAA 180
181 TCTGCAAGCAATTTCCCTCCAAAGTGGATGATCACTCTCACCAACAGTTGATTAAGTTGAC 240
181 TCTGCAAGCAATTTCCCTCCAAAGTGGATGATCACTCTCACCAACAGTTGATTAAGTTGAC 240
241 TTAACCTTTCAACTTCGACCTTCGACAGCACTATTTCTGTAAGGCTTCAGGAAGCTTGGTT 300
241 TTAACCTTTCAACTTCGACCTTCGACAGCACTATTTCTGTAAGGCTTCAGGAAGCTTGGTT 300
301 ACCGATTCAGTTAGAACCAAAAGCAACCAAGCAAAATCTTGACTACTATTAATAATA 360
301 ACCGATTCAGTTAGAACCAAAAGCAACCAAGCAAAATCTTGACTACTATTAATAATA 360
361 GGTCTCATGTCTAATGGGAACTATGGAATGACAGAGCAAGCTCAAAAGCTACTATGTTAA 420
361 GGTCTCATGTCTAATGGGAACTATGGAATGACAGAGCAAGCTCAAAAGCTACTATGTTAA 420
421 AGACCTCAATTAATTAAGTTTACCTGATTAAGCTTGTGCTGGAATGCTTACAGCACC 480
421 AGACCTCAATTAATTAAGTTTACCTGATTAAGCTTGTGCTGGAATGCTTACAGCACC 480
481 AAACCAATATGACCCCTATTCACATCCAGAGAGAGCCCAAGAGCCGCCAACTTGGCTTT 540
481 AAACCAATATGACCCCTATTCACATCCAGAGAGAGCCCAAGAGCCGCCAACTTGGCTTT 540
541 ATCTGAATATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAAAGCACTAATAC 600
541 ATCTGAATATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAAAGCACTAATAC 600
601 ACCAATTAATGATGACCTCAAAAGTCTCAATATCAGCAAGTAATTAATCCCTGCTTACATGGA 660
601 ACCAATTAATGATGACCTCAAAAGTCTCAATATCAGCAAGTAATTAATCCCTGCTTACATGGA 660
661 TTAATTAACCTCAAGAGATCATCAATCAAGTGAAGAAAGCAAGGCTATTAACCTACTCTAC 720
661 TTAATTAACCTCAAGAGATCATCAATCAAGTGAAGAAAGCAAGGCTATTAACCTACTCTAC 720
721 AACTGGGATGATGATCTTACCAAAATGTAGAACCAAGAGCTCAAAACATCTGTGGATAT 780
721 AACTGGGATGATGATCTTACCAAAATGTAGAACCAAGAGCTCAAAACATCTGTGGATAT 780
781 TTACATATACAGAGAAATACGTTGCTATCCAGATGATGATGATGATGATGATGATGAT 840
781 TTACATATACAGAGAAATACGTTGCTATCCAGATGATGATGATGATGATGATGATGAT 840
841 TGTGTATGTTTCTTAAGGTTAAGTCTATTTGCCAGCTAGAGACAGCCATCATGCAAGTAA 900
841 TGTGTATGTTTCTTAAGGTTAAGTCTATTTGCCAGCTAGAGACAGCCATCATGCAAGTAA 900
901 TGTTCCTCTCGAATTAACCAAGCAAGTGAAGAAACCCGAGCTGGGGATCAACTATGAA 960
901 TGTTCCTCTCGAATTAACCAAGCAAGTGAAGAAACCCGAGCTGGGGATCAACTATGAA 960
961 ACCGATACAGACTATGCTGCTGCTTGGAGTACGGTGTCTACGATTCAGTCTACTAT 1020
961 ACCGATACAGACTATGCTGCTGCTTGGAGTACGGTGTCTACGATTCAGTCTACTAT 1020
1021 CGTTACAGATGAGCCCTATTAATCACTCCCTGGGACAAATTAATTAATTAATTAATTAAT 1080
1021 CGTTACAGATGAGCCCTATTAATCACTCCCTGGGACAAATTAATTAATTAATTAATTAAT 1080
1081 GGGCTACTTGGCAACATACCTTGCATATAGCCCTGCAACAATGCGGAACGCTCCACG 1140
1081 GGGCTACTTGGCAACATACCTTGCATATAGCCCTGCAACAATGCGGAACGCTCCACG 1140

1141 CGTGAAGAACTTAACAGAGCTGGACTCAACCGCGCCAAACACTTCTTAATAGTCTAGG 1200
1141 CGTGAAGAACTTAACAGAGCTGGACTCAACCGCGCCAAACACTTCTTAATAGTCTAGG 1200
1201 AATGCACTACCCCAAGTATTCACACTACATCAATGGCATTTTCAAGTAAACCAACCGCAATCAGA 1260
1201 AATGCACTACCCCAAGTATTCACACTACATCAATGGCATTTTCAAGTAAACCAACCGCAATCAGA 1260
1261 CAAAAATATGAGCAAGTATGTAAGAAAGATGCTGCTGCTTACGCTTCCCTTCAAAATGG 1320
1261 CAAAAATATGAGCAAGTATGTAAGAAAGATGCTGCTGCTTACGCTTCCCTTCAAAATGG 1320
1321 TGGAACTTACTATTAACCAATGATATATCAATAAAGTCTGCTTATGATGAGGAGTAAAA 1380
1321 TGGAACTTACTATTAACCAATGATATATCAATAAAGTCTGCTTATGATGAGGAGTAAAA 1380
1381 AGAGTTCTCTAATGTGGAACTGTGCTGATGAAGGAAAGCAGACGCTATATATGATGACCA 1440
1381 AGAGTTCTCTAATGTGGAACTGTGCTGATGAAGGAAAGCAGACGCTATATATGATGACCA 1440
1441 CATGATGAAAAACAGTCTTGTGACTATGTAAGTGAACCTGAGCAAAATGCTATCTTGGCTGCC 1500
1441 CATGATGAAAAACAGTCTTGTGACTATGTAAGTGAACCTGAGCAAAATGCTATCTTGGCTGCC 1500
1501 TCAGGCTGTGTAACAGAGAACTCTAATATACAGACAGAGAAATTTGAAAAACCAATCAA 1560
1501 TCAGGCTGTGTAACAGAGAACTCTAATATACAGACAGAGAAATTTGAAAAACCAATCAA 1560
1561 GACCTCTCAATTTGTAGCAACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1561 GACCTCTCAATTTGTAGCAACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1621 GGTGTATGAGAGAGGCTATTTCTTAACCGTGTGACACCACTTGTGAGGAATGAGCTTACGGT 1680
1621 GGTGTATGAGAGAGGCTATTTCTTAACCGTGTGACACCACTTGTGAGGAATGAGCTTACGGT 1680
1681 CGCTGCAAAAGTTTACCGCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1681 CGCTGCAAAAGTTTACCGCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1741 TTGGAAATATACAGAGAGGCTCTACAGAAATGGAAGAAATTCATTTAAAAATGAGTCTCG 1800
1741 TTGGAAATATACAGAGAGGCTCTACAGAAATGGAAGAAATTCATTTAAAAATGAGTCTCG 1800
1801 TTCTACGTGAATCTACGCTGCTCCACAAACACCCCATCAACCTGAAGTTCAAGCTCATC 1860
1801 TTCTACGTGAATCTACGCTGCTCCACAAACACCCCATCAACCTGAAGTTCAAGCTCATC 1860
1861 ATCAGATAGTTCAACTTCAACAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1861 ATCAGATAGTTCAACTTCAACAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1921 TACCAATCTTAACCAATTAATGCAACCAATTAATTAATTAATTAATTAATTAATTAAT 1980
1921 TACCAATCTTAACCAATTAATGCAACCAATTAATTAATTAATTAATTAATTAATTAAT 1980
1981 TCCCTAACCAAGCAACCA 1999
1981 TCCCTAACCAAGCAACCA 1999

RESULT 2
US-08-481-435-5
Sequence 5. Application us/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balganes, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas

```
? CITY: New York
? STATE: New York
? COUNTRY: United States
? ZIP: 10036-2787
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/481,435
? FILING DATE: 10-JUL-1995
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: IN 580/MAS/94
? FILING DATE: 01-JUL-1994
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: SE 9404072-2
? FILING DATE: 24-NOV-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Sterner, Richard J.
? REGISTRATION NUMBER: 35,372
? REFERENCE/DOCKET NUMBER: 1103326-151
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 819-8783
? TELEFAX: (212) 354-8113
? INFORMATION FOR SEQ. ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2049 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? ORIGINAL SOURCE:
? ORGANISM: Streptococcus pneumoniae
? STRAIN: PM 1
? IMMEDIATE SOURCE:
? LIBRARY: PCR cloning
? CLONE: PARC 0512 Soluble PBP 1A del 38
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..2049
? FEATURE:
? NAME/KEY: mat.peptide
? LOCATION: 1..2046
? US-08-481-435-5

Query Match          99.1%; Score 1981.4; DB 3; Length 2049;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```
|||||
Db 348 AGCGATTCACTTGAACAAAGCAACCAAGCAAGAAATCTTGACACTATATATAATAA 407
Qy 361 GGTCACATGTCTATATGGGAACCTATGGAATGCAGACAGCGCTCAAACTACTATGTAA 420
Db 408 GGCTACATCTCTAATGGGAACCTATGGAATGCAGACAGCGCTCAAACTACTATGTAA 467
Qy 421 AGACCTCAATATTAATTAAGTTTACCTCGATTAGCCTTGTGCTGGAAATGCTCAGGCACC 480
Db 468 AGACCTCAATATTAATTAAGTTTACCTCGATTAGCCTTGTGCTGGAAATGCTCAGGCACC 527
Qy 481 AAACCAATATGACCCCTATTACATCCAGAGAGAGCCCAAGCCGCGAAACTTGGCTT 540
Db 528 AAACCAATATGACCCCTATTACATCCAGAGAGAGCCCAAGCCGCGAAACTTGGCTT 587
Qy 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCACTCAATAC 600
Db 588 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCACTCAATAC 647
Qy 601 ACCAATTACGTAGGACTACAAAGCTCAATGACGAAGTAATTAACCTGCTTACATGA 660
Db 648 ACCAATTACGTAGGACTACAAAGCTCAATGACGAAGTAATTAACCTGCTTACATGA 707
Qy 661 TAAATTACCTCAAGGAGTCAATCAATCAAGTTGAAGAAAGCAAGGCTAATACCTACTAC 720
Db 708 TAATTACCTCAAGGAGTCAATCAATCAAGTTGAAGAAAGCAAGGCTAATACCTACTAC 767
Qy 721 AACTGGAGTATGCTCTACCAAAATGTAGACCAAGAAAGCTCAAAACATCTGTGGATAT 780
Db 768 AACTGGAGTATGCTCTACCAAAATGTAGACCAAGAAAGCTCAAAACATCTGTGGATAT 827
Qy 781 TTACAATACAGAGAAATACGTTGCTTCCAGAGAAATGCAATGCAATGCTTCTTACAT 840
Db 828 TTACAATACAGAGAAATACGTTGCTTCCAGAGAAATGCAATGCAATGCTTCTTACAT 887
Qy 841 TGTGATGTTTCTAACGGTAAAGTCAATGCTGCCAGCTAGAGACAGCCATCAGTCAAGTAA 900
Db 888 TGTGATGTTTCTAACGGTAAAGTCAATGCTGCCAGCTAGAGACAGCCATCAGTCAAGTAA 947
Qy 901 TGTTCCTTGGGAATTAACCAAGCAAGTGAACCAACCGGAGCTGGGATCAACTATGAA 960
Db 948 TGTTCCTTGGGAATTAACCAAGCAAGTGAACCAACCGGAGCTGGGATCAACTATGAA 1007
Qy 961 ACCGATCACAGACTATGCTCTGCTGAGTACGGTGTCTAGGATTAACCTGCTACTAT 1020
Db 1008 ACCGATCACAGACTATGCTCTGCTGAGTACGGTGTCTAGGATTAACCTGCTACTAT 1067
Qy 1021 CGTTACAGATGAGCCCTATTAACCTGCGGACAAATACCTGTTATTAACCTGGATAG 1080
Db 1068 CGTTACAGATGAGCCCTATTAACCTGCGGACAAATACCTGTTATTAACCTGGATAG 1127
Qy 1081 GGGCTACTTTGGCAACATCACTTGCAATAGCCCTTGCAAAATCGGAAACGTCCAGC 1140
Db 1128 GGGCTACTTTGGCAACATCACTTGCAATAGCCCTTGCAAAATCGGAAACGTCCAGC 1187
Qy 1141 CGTGGAAACTTAACCAAGGTGGAGTCAACCGCGCCCAACACTTTCTTAATAGTCTAGG 1200
Db 1188 CGTGGAAACTTAACCAAGGTGGAGTCAACCGCGCCCAACACTTTCTTAATAGTCTAGG 1247
Qy 1201 AATGACTACCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
Db 1248 AATGACTACCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1307
Qy 1261 CAAAAAATATGAGCAAGTATGAGAAAGATGGCTGCTGTACAGCTGCTTTCGAAATGG 1320
Db 1308 CAAAAAATATGAGCAAGTATGAGAAAGATGGCTGCTGTACAGCTGCTTTCGAAATGG 1367
Qy 1321 TGAACCTTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
Db 1368 TGAACCTTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1427
Qy 1381 AGAGTTCTTAATGTCGGAAGTCTGTCATGAGGAAGCAACAGCCATATATGATGACCGA 1440
```



```

Db      612  ACCAATTACGATGGGCTACCAAGTCTCAAAATCAGCAAGTAATTAACCTTGCTTACATGGA 671
Oy      661  TAAATTACCTCAAGGAAGTCATCAATCAAGTTGAGAAGAAACAGCGTATTAACCTACTACAC 720
Db      672  TAAATTACCTCAAGGAAGTCATCAATCAAGTTGAGAAGAAACAGCGTATTAACCTACTACAC 721
Oy      721  AACTGGGATGGATGTCTACACAAATGTAGACCAAGAAGCTCAAAAACATGTGGGATAT 780
Db      732  AACTGGGATGGATGTCTACACAAATGTAGACCAAGAAGCTCAAAAACATGTGGGATAT 791
Oy      781  TTCAATATACAGACGAATACGTTGGCTATCCAGACGATGATTCGAATGCGCTCTACCAT 840
Db      792  TTCAATATACAGACGAATACGTTGGCTATCCAGACGATGATTCGAATGCGCTCTACCAT 851
Oy      841  TGTTCGATGTTCTTAACGCTAAAGTCATTTGCCACGCTAGAGACAGCCATCAGTCAAGTAA 900
Db      852  TGTTCGATGTTCTTAACGCTAAAGTCATTTGCCACGCTAGAGACAGCCATCAGTCAAGTAA 911
Oy      901  TGTTCCTTCCGGAATTTAACCAAGCAGTAGAATAACACCGCGACTGGGGA 949
Db      912  TGTTCCTTCCGGAATTTAACCAAGCAGTAGAATAACACCGCGACTGGGGA 960

RESULT      4
US-08-600-993A-3
; Sequence 3, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Measure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauder & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 960 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)

```

	Query Match	47.2%	Score 944.2;	DB 2;	Length 960;	
	Best Local Similarity	99.7%;	Pred. No. 1.9e-278;			
	Matches 946;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
Y	1 TAAATCTACGCATTAATAAATCAACCTGATGCTGCTGTTGGTTGGAAGCGCGCTCAA	60				
Db	12 TAAATCTACGCATTAATAAATCAACCTGATGCTGCTGTTGGTTGGAAGCGCGCTCAA	71				
Y	61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACA	120				
Db	72 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACA	131				
Y	121 TCGCTTCTTCGACCCAGGGGGATTTGATACCATCCGTAATCCTGGGAGCTTCTTCGGCAA	180				
Db	132 TCGCTTCTTCGACCCAGGGGGATTTGATACCATCCGTAATCCTGGGAGCTTCTTCGGCAA	191				
Y	181 TCTGCAAGCAATTCCTCCACAGGGGATGATCACTCCACCAAGTATTAGTTAGTTGAC	240				
Db	192 TCTGCAAGCAATTCCTCCACAGGGGATGATCACTCCACCAAGTATTAGTTAGTTGAC	251				
Y	241 TTACTTTTCAACTTGACCTCGACACAGTATTTCTGTAAGGCTCAGGAAGCTTGGTT	300				
Db	252 TTACTTTTCAACTTGACCTCGACACAGTATTTCTGTAAGGCTCAGGAAGCTTGGTT	311				
Y	301 AGCGATTGATTTAGAACAAAAAGCACCAGACAGAAATCTTGACCTCTATATATAATA	360				
Db	312 AGCGATTGATTTAGAACAAAAAGCACCAGACAGAAATCTTGACCTCTATATATAATA	371				
Y	361 GGTTCACATGCTCTATGGAAGAACTATGGAATGAGACAGACAGTCAAACCTATAGTAA	420				
Db	372 GGTTCACATGCTCTATGGAAGAACTATGGAATGAGACAGACAGTCAAACCTATAGTAA	431				
Y	421 AGACCTCAATATATTTAAGTTTACCTGATGACCTTGGCTGGCTGGGAATGCTCAGGACCC	480				
Db	432 AGACCTCAATATATTTAAGTTTACCTGATGACCTTGGCTGGGAATGCTCAGGACCC	491				
Y	481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGCCGCGCAAACTTGGTCTT	540				
Db	492 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGCCGCGCAAACTTGGTCTT	551				
Y	541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGGAAGAAGCAGTCAATAC	600				
Db	552 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGGAAGAAGCAGTCAATAC	611				
Y	601 ACCAATTAAGTATGAGCTACAAAAGTCTCAAAATCAGCAAGTAATTAACCTCGCTTACATGA	660				
Db	612 ACCAATTAAGTATGAGGCTACAAAAGTCTCAAAATCAGCAAGTAATTAACCTCGCTTACATGA	671				
Y	661 TAAATTAAGTCAAGGAAGCATCAATCAAGTGAAGAAAGAAAGAGCTTAACCTATAC	720				
Db	672 TAAATTAAGTCAAGGAAGCATCAATCAAGTGAAGAAAGAAAGAGCTTAACCTATAC	731				
Y	721 AACTGGATGATGCTCTACACAAATGTTAGACCAAGAACTCAAAAACATCTGTGGGATAT	780				
Db	732 AACTGGATGATGCTCTACACAAATGTTAGACCAAGAACTCAAAAACATCTGTGGGATAT	791				
Y	781 TTACATATACAGACGAATATGTTGCTTATCAAGAGATGATTTGCAAGTGTCTTACCAT	840				
Db	792 TTACATATACAGACGAATATGTTGCTTATCAAGAGATGATTTGCAAGTGTCTTACCAT	851				

Db 200 GCTTCTACGACATACAGCGCGTTGACCCTGGGGATCTTCCTGACAGCAAGCGTGCGC 259
OY 183 TGCAGCAATTCCTCCCTCAAGGTGATCACTCTACCCCAAGTTGATTAAGTACTT 242
Db 260 TGTTCCTCGGTCACGCGTCACAAAGGGCAAGTACCATTACCAGAGCTGGCGAGAACT 319
OY 243 ACTTTCAACTTCGACTTCGACAGCAAGTATTTCTCGTAGGCTCAGGAAGCTTGGTTAG 302
Db 320 TCTTCCTACAGTCC-----AGAACGACGCTGATGCCGTAAAGATTAAAGAACTCTTCG 373
OY 303 CGATTCACTTGAACAAAAGCAACCAAGCAAGAAATCTTACCTACTATTAATAAG 362
Db 374 CGATTCCATGGAACAGCTGCTGACGAAAGAGATCCTGAGAGTTATCTGAAACAAGA 433
OY 363 TCTCATGTCTAATGGGAGCTATGAAATGCAGACAGCAGCTCAAAACTACTAGTGAAG 422
Db 434 TTTCCTTGGTTACCGCGCTATGTGTGCTGCTGCGGCAAGTCTATTTCGAAATA 493
OY 423 ACCTCAATTAATTTAATTTTACCTCAATTAAGCTTGTGCTGGAATGCTCAGGACCA 482
Db 494 CGGTGACCACTGACGCTGAACGAAATGGGCGTATAGCGCGGCTGCCGAAGGCGCTT 553
OY 483 ACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGAACTTGGTCTTAT 542
Db 554 CCACCTTCAACCCCTCTACTCTGATGATCGTCCGCTGCGCGGCTAACCTGCTGT 613
OY 543 CTGAATGAATAATCAAGGCTACATCTGCTGAACAGTATGAGAAACACTCAATAC 602
Db 614 CGCGAATGCTGATGAAAGGATATACCCCAACAGCTGATGACAGACGACTGAGG 673
OY 603 CAATTAAGT 613
Db 674 CGATTAAGCT 684

RESULT 8
US-08-771-716-1
Sequence 1, Application US/08771716
Patent No. 5922540
GENERAL INFORMATION:
APPLICANT: Peery, Robert B.
APPLICANT: Hoskins, Joann
APPLICANT: Jaskunas, S. Richard
APPLICANT: Skatrud, Paul L.
TITLE OF INVENTION: Monofunctional Glycosyltransferase
TITLE OF INVENTION: Gene of Staphylococcus Aureus
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/771.716
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11067
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 807 base pairs
type: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..807
US-08-771-716-1

Query Match 3.18; Score 62.2; DB 2; Length 807;
Best Local Similarity 49.98; Pred. No. 3.4e-09;
Matches 191; Conservative 0; Mismatches 183; Indels 9; Gaps 1;

Db 67 AGCTAATATATTTCCACAGATTGTTAAGGCATCTTCTCTACGAAAGCAATCGCTT 126
OY 252 AGCTGATTAACATGCCAGATATGTTAAAGTGCTTATTTCAATGGAAGTGAACGATT 311
Db 127 CTTCGACCACAGGGGATGATACCATCCGATTCCTGGAGCTTCTTCCCAATCTCA 186
OY 312 CTACATATCATGATTCGATTTGAAAGTACACTAGACCTTATTTTCACAGATTAG 371
Db 187 AAGCAATTCCTCCCAAGGTGATCACTCTCACCAAGTATTAAGTGAATGACTTACTT 246
OY 372 CGACAGATGTCGAAGGTGATGATCAATTAACAACAGTTGCAAAATTAATTTTA 431
Db 247 TTCAACTTCGACTTCGACGACGACTATTTTCGCTGAAGCTCAGAGAGCTTGTAGCAT 306
OY 432 TGATTAAT-----GATTCCTTCAATTTCTAGAAAGTAAAGATTAATTTTGTCTCA 482
Db 307 TCAGTTAGACAAAAGCAACCAAGCAAGAAATCTTCACTACTATTAATAAGTCTA 366
OY 483 TCGAGTTTAAACATATTAATTAAGAACAAATTTTAAGCTTTTAAATTAATTAATTTA 542
Db 367 CATGCTAATGGAATGGAATGCAAGACAGACAGCTCAAAACTACTATTAAGTAAAGACT 426
OY 543 CTTGGCGATTAATCAATTAATGCTTGAAGGCGCAGCAACCATTAATTTGGAACACCGT 602
Db 427 CAATTAATTAATTTTACTCTAGT 449
OY 603 GAATTAATAATGATCAACCAATGT 625

RESULT 9
US-08-771-716-3
Sequence 3, Application US/08771716
Patent No. 5922540
GENERAL INFORMATION:
APPLICANT: Peery, Robert B.
APPLICANT: Hoskins, Joann
APPLICANT: Jaskunas, S. Richard
APPLICANT: Skatrud, Paul L.
TITLE OF INVENTION: Monofunctional Glycosyltransferase
TITLE OF INVENTION: Gene of Staphylococcus Aureus
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/771.716
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

```

1 RESULT 10
2 US-09-057-720A-1
3 : Sequence 1, Application US/09057720A
4 : Patent No. 6143868
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Peery, Robert B.
9 :
10 : APPLICANT: Hoskins, Joann
11 :
12 : APPLICANT: Jaskunas, S. Richard
13 :
14 : APPLICANT: Skatrud, Paul L.
15 :
16 : TITLE OF INVENTION: Monofunctional Glycosyltransferase Gene
17 :
18 : TITLE OF INVENTION: of Staphylococcus Aureus
19 :
20 : NUMBER OF SEQUENCES: 3
21 :
22 : CORRESPONDENCE ADDRESS:
23 :
24 : ADDRESSEE: Eli Lilly and Company
25 :
26 : STREET: Lilly Corporate Center
27 :
28 : CITY: Indianapolis
29 :
30 : STATE: Indiana
31 :
32 : COUNTRY: U.S.
33 :
34 : ZIP: 46285
35 :
36 : COMPUTER READABLE FORM:
37 :
38 : MEDIUM TYPE: Floppy disk
39 :
40 : COMPUTER: IBM PC compatible
41 :
42 : OPERATING SYSTEM: PC-DOS/MS-DOS
43 :

```

RESULT 11
US-09-057-720A-3
Sequence 3, Application US/09057720A
Patent No. 6143868
GENERAL INFORMATION:
APPLICANT: Peery, Robert B.
APPLICANT: Hoskins, JoAnn
APPLICANT: Jaskunas, S. Richard
APPLICANT: Skatrud, Paul L.
TITLE OF INVENTION: Monofunctional Glycosyltransferase Gene
TITLE OF INVENTION: of Staphylococcus Aureus
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company

RESULT 12
US-08-731-716-1
Sequence 1, Application US/08731716
Patent No. 5789202
GENERAL INFORMATION:
APPLICANT: Hoskins, Joann
APPLICANT: Jackson, S. Richard
APPLICANT: Rockey, Pamela K.

[illegible]

```

Dbb 625 TATTTTGGAAATGCTGTGTGGGGGTGTGAGAAATGCGAGTAAGAAATACTTTGGAGTTTCT 664
Qy 425 CTCATAATTAATTAAGTTTAACTTCACTCACTTACAGCTTGTGCTGGCTGGAATGCCCTACGCCAAC 484
Db 685 GCATCAGAGAAAGTATCTGGATTCGAAGCTGCACACTCTTGGCAGGGATGCTCAAGGGGGCGGAA 744
Qy 485 CAATATGACCCCTTTCACATCCGAGAGCAGGCCAAGACCGCGGAAACTGGTCTATCTG 544
Db 745 CTGTATTAATCCCTTGAATTCCTGTAAGAAAGATTTCTACTAATATGCGCGGATTACTGTTTGCAG 804
Qy 545 GAAATGAAAAAATCAAGGCTCAATCTCGCTGGAACAGATATGAGAAAGCAGTCAATATACCA 604
Db 805 AATATGTTGTCAGCAGGATATATATGATTAATAAACCAAAACCAAGCTGCTGAAGTTGAT 864
Qy 605 ATTAAGTGAAGCTACAAAGCTCTCAAAATCAGCAAGTATTAATACCTCGCTTACATGATTAAT 664
Db 865 ATGACTTCGCAATTCGACAGTAAAGTATGAAGAAATAATCTCAAGATTACCTTACCCCTCT 924
Qy 665 TAACTCAAGAAAGTCATCATCAAGTTG-----AAAGAAACAGGCTTAAACCTA 715
Db 925 TATTTTGAATGGGGTGAATTAAGAGCTGTTTCCAAAGTATATATCAACAGAGGAAGAGATT 984
Qy 716 CTCACAACTGGAGATGATGTCTTACACAAATGTATGACCAAGACAGCTCAAAAATCATGTGG 775
Db 985 GTCAATATATGCTTACCCGATTTTACACAGAGCTGGACCAAAATCAACAGCAAAATATGCA 104
Qy 776 GATATTTTACATACAGCAGAAATAGCTTGCCCTATCCAGACAGATGAATTCAGAGTGGCTTC 835
Db 1045 ATGTGTTTATGAAGAACATATGCTATTTCCGAGGGACAGATGGAAGCTTGTGCTCAATCA 110
Qy 836 ACCATTGTTGATGT 849
Db 1105 GGAAGTGTAGCTCT 1118

RESULT 13
US-08-731-716-3
; Sequence 3, Application US/08731716
; Patent No. 5789202
; GENERAL INFORMATION:
; APPLICANT: Hoskins, Joann
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Rocky, Pamela K.
; APPLICANT: Zhao, Genshi
; APPLICANT: Rostek, Paul R. Jr.
; APPLICANT: No. 5789202, Frank H.
; TITLE OF INVENTION: Penicillin Binding Protein From
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,716
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-10,887
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

	LENGTH: 2193 base pairs	
	TYPE: nucleic acid	
	STRANDEDNESS: single	
	TOPOLOGY: linear	
	MOLECULE TYPE: mRNA	
	HYPOTHETICAL: NO	
	ANTI-SENSE: NO	
	US-08-731-716-3	
Query Match	2.8%; Score 56.4; DB 1; Length 2193;	
Best Local Similarity	31.6%; Pred. No. 3.4e-07;	
Matches	270; Conservative 111; Mismatches 446; Indels 27; Gaps	
OY	5 ATCTAGACATAATAAATCACTCATCTGATCTGACTTGGTTCGACGCCGCTCAATGCC 64	
Db	283 AATUUUUAGCCGUGAAGAAAAAGAGCGUGGUCUUUGUCUCAAAGGACCAUAGUU 342	
OY	65 CAACCTATGATATATCCACAGATTTGGTTAAGGCAATGCTTTCTATGCAAAACCATGCG 124	
Db	343 GAGCGUGACUGCAACAGUAAAAACUGGCAGAAAGCGUUAUUGCGACAGAAACCGUUCU 402	
OY	125 TTCTTCACACCAAGGCGATTGGATACCAACCGATCCGTGGAGCTTTCTTGGCAATCTG 184	
Db	403 UUCUUAUAAAAUAGCGGGAUUAACUADGGCCGUUC-----UUCUUGCUAUUGUC 453	
OY	185 CAAAGCAATTTCCCTCCAAAGGTGGATCAACTCTCCACCAAGTGTAAATTGACTTAC 244	
Db	454 ACUCUGGACGUCUADGUGUGUGUCUCUACCUUATCCCAACAGUGUGCUAAAAAGCCUAG 513	
OY	245 TTTTCAACTTCGACTTCGACACGACTATTCTCGTAAGGCTCAGGAAGCTTGGTTAGCG 304	
Db	514 UUAU-----CGCAGGAUCAAACUGUUGAGAGAAAAAGCAAGAAUUUUUCCUUGCC 564	
OY	305 ATTGAGTTAGAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATTAATTAAGGTC 364	
Db	565 UUGAUAUUAGCAAAAAAUUAGUUAAGGACCAAAUUCUAAACUAGACCUUAACACGU 624	
OY	365 TACATGTCTATGAGGAACCTATGGAATGACAGACAGCACTCAAAACTCATGTGTAAGAC 424	
Db	625 UAUUUUGGAAAUUGUGUGUGGUGUGUAAGAAUUGCAGUAAGAAUACUUGUGAGUUCU 684	
OY	425 CTCATTAATTTAACTTAACTCACTGACGTTAGCCTTCTGCTGGAATGCTCAGCACCAC 484	
Db	685 GCAUCACAAAGUAGUUCUGUAGUCAAAGCUGCAGUCGACGAGGAGUUGCUAAGGGCCGAA 744	
OY	485 CAATATGACCCCTTATTCACATCCAGAAAGCAGCCCAAGCCGCGAATCTTGCTTATCT 544	
Db	745 CUGAUAUUCCCUUGAUAUUCGUGAAGAUUUCUAUAUCCGGCGCAUUCUUCUUGCAG 804	
OY	545 GAATATGAAAAATCAAGGCTCATCTCGTCGAGACATATGAGAAAGCACTCAATACACA 604	
Db	805 AAUUAUGGUGCAGCAGAGAUUUAUUGAUAAAAAACCAAGAAACCGAAGCUCUGAAUUGAU 864	
OY	605 ATTACTGATGACTACCAAGGTCTCAATCAGCAAGTAATTACCTGCTTACATGATATAT 664	
Db	865 AUGACUUCGCAUUDGACGAAUAGUAUAGAAAGAAAUUCUAGAUUACCGUUAACCCUUCU 924	
OY	665 TACTCTCAAGGAAGTCATCATCAAGTTG-----AAGAGAAACAGGCTATTAACCTA 715	
Db	925 UAUUUUADUGCGGUGUUAUAGGUCUGUUCCAAUUAUAUCUAAACAGAGAAAGAUU 984	
OY	716 CTCACAACTGGGATGATGTGCTACACAATATGTAGACCAAGAAAGCTCAAAAACATGTCG 775	
Db	985 GUCUAUAUUGCUACCGCAUUDUACACGAGCUGGACACCAAAUACCAAGCAAAUUDGCA 1044	
OY	776 GATATTTTACATACAGACGAATAGTTCCTATCTCAGACGATGAATTTGCAAGTCTGCT 835	
Db	1045 AATUUUUUAGAAAAACAGAVCGUAUUUCCGAGGCGCAGAGAGUAGAAACUUUCCUACUA 1104	
OY	836 ACCATTGTTGATGT 849	
Db	1105 GGAAGUGUAGUCUU 1118	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2002, 01:34:34 ; Search time 3439.02 Seconds
(without alignments)
12575.723 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTGACGACATATAA.....ATTCCTACACGACACACCA 1999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/PCRUS.COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06.COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07.COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080.COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081.COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082.COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083.COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084.COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085.COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086.COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087.COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088.COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089.COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090.COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091.COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092.COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093.COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094.COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095A.COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095B.COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095C.COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095D.COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US096A.COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096B.COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096C.COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096D.COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096E.COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US097A.COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097B.COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097C.COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US098A.COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US098B.COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US098C.COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US099A.COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US099B.COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US099C.COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US100.COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US101.COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US6000.COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US6001.COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US6002.COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US6003.COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US6004.COMB.seq:*

44: /cgn2_6/ptodata/2/pna/US6005.COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US6006.COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US6007.COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US6008.COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US6009.COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US6010.COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US6011.COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US6012.COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US6013.COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US6014.COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US6015.COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US6016.COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US6017.COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US6018.COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6019.COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6020.COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6021.COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6022.COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6023.COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6024.COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6025.COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6026.COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6027.COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6028.COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6029.COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6030.COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6031.COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6032.COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6033.COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6034.COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6035.COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6036.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	20	US-09-536-784-1
2	1999	100.0	1999	30	US-08-765-271-1
3	1999	100.0	1999	30	US-09-765-272-1
4	1999	100.0	10383	41	US-60-029-960-81
5	1999	100.0	10711	13	US-08-961-527-145
6	1987.8	99.4	2160	22	US-09-583-110-1312
7	1987.8	99.4	2166	15	US-09-107-433-1102
8	1981.4	99.1	2160	11	PCT-US02-03987-9325
9	1981.4	99.1	2160	31	US-09-815-242-9325
10	1981.4	99.1	2160	37	US-10-072-851-9325
11	1981.4	99.1	10333	45	US-60-061-998-596
12	1631.2	81.6	9845	45	US-60-068-175-596
13	944.2	47.2	960	5	US-08-116-541-3
14	808.6	40.5	2172	1	PCT-US97-14436-139
15	808.6	40.5	2172	13	US-08-911-503-139
16	808.6	40.5	2172	13	US-08-911-503A-139
17	574.6	28.7	2337	1	PCT-US02-03987-6631
18	574.6	28.7	2337	31	US-09-815-242-6631
19	574.6	28.7	2337	37	US-10-072-851-6631
20	573	28.7	2370	15	US-09-134-000-1534
21	559	28.0	8395	14	US-09-070-927-217
22	559	28.0	8395	14	US-09-070-927A-217
23	553	27.7	2472	15	US-09-107-532-2013
24	553	27.7	2472	15	US-09-107-532A-2013
25	496.2	24.8	1386	37	US-10-091-007-91
26	410	20.5	2470	44	US-60-068-186-495
27	410	20.5	2474	44	US-60-050-444-495
28	352.2	17.6	3478	44	US-60-068-217-710
29	348.6	17.4	3279	25	US-09-634-238-44
30	348.6	17.4	7818	24	US-09-634-238-14
31	273	13.7	933	43	US-60-045-649-108

0Y	1201	AATGACATACCAAGATTTCATACACAAATGCAATTCATAATACACACCGAATCGA	1260
0Y	1201	AATGACATACCAAGATTTCATACACAAATGCAATTCATAATACACACCGAATCGA	1260
Db	1201	AATGACATACCAAGATTTCATACACAAATGCAATTCATAATACACACCGAATCGA	1260
0Y	1261	CAAAAAATATGAGACAGTAGTGAAGAAAGATGGCTGCTTACGCTGGCTTTGCAAAATG	1320
Db	1261	CAAAAAATATGAGACAGTAGTGAAGAAAGATGGCTGCTTACGCTGGCTTTGCAAAATG	1320
0Y	1321	TGGAACTTACTATTAACCAATGATATTCATAAAGTCGCTTTAGTATGGAGTGA	1380
Db	1321	TGGAACTTACTATTAACCAATGATATTCATAAAGTCGCTTTAGTATGGAGTGA	1380
0Y	1381	AGAACTTCATAAATGTGGAACTCTGTCGCATGGAAGAAAGACAGGCTATATGATGACGA	1440
Db	1381	AGAACTTCATAAATGTGGAACTCTGTCGCATGGAAGAAAGACAGGCTATATGATGACGA	1440
0Y	1441	CATGATGAAAAACAGTCTTGACTTATGGAACCTGGAGCAATGGCTATCTTGCTGGCTCC	1500
Db	1441	CATGATGAAAAACAGTCTTGACTTATGGAACCTGGAGCAATGGCTATCTTGCTGGCTCC	1500
0Y	1501	TCAGAGCTGGTAAAAACAGGAACCTCTTAATATACAGACGAGAAATTGAAAAACACATCAA	1560
Db	1501	TCAGAGCTGGTAAAAACAGGAACCTCTTAATATACAGACGAGAAATTGAAAAACACATCAA	1560
0Y	1561	GACCTCCATATTTGTATGACACCTATATACCTATTTTGTGGCTTACGCTAAATATTCAT	1620
Db	1561	GACCTCCATATTTGTATGACACCTATATACCTATTTTGTGGCTTACGCTAAATATTCAT	1620
0Y	1621	GGCTGTATGACAGGCAATTCATACGCTGACACACTTTGAGGCATAGGCTTACGCT	1680
Db	1621	GGCTGTATGACAGGCAATTCATACGCTGACACACTTTGAGGCATAGGCTTACGCT	1680
0Y	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTAACCTGTCTGAAGAGCAATCCAGAAGA	1740
Db	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTAACCTGTCTGAAGAGCAATCCAGAAGA	1740
0Y	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAAAATGGTCTCG	1800
Db	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAAAATGGTCTCG	1800
0Y	1801	TTTACGCTGGAACCTCACCTCTCCACAAACACCCCATCACTGAAGCTCAAGCTCATC	1860
Db	1801	TTTACGCTGGAACCTCACCTCTCCACAAACACCCCATCACTGAAGCTCAAGCTCATC	1860
0Y	1861	ATCAGATAGTCAACTTCACAGCTCTAGCTCAACCACTCCACACACAAATTAATAGTAGAC	1920
Db	1861	ATCAGATAGTCAACTTCACAGCTCTAGCTCAACCACTCCACACACAAATTAATAGTAGAC	1920
0Y	1921	TACCAATCTTACCAATATATACGACAAATCAATACACCCCTGATCAACAAATTCAGAA	1980
Db	1921	TACCAATCTTACCAATATATACGACAAATCAATACACCCCTGATCAACAAATTCAGAA	1980
0Y	1981	TTCTTCACACGACGACAAACCA 1999	
Db	1981	TTCTTCACACGACGACAAACCA 1999	
RESULTS			
US-09-765-271-1			
Sequence 1, Application US/09765271			
GENERAL INFORMATION:			
APPLICANT: Choi et. al.			
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines			
NUMBER OF SEQUENCES: 452			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Human Genome Sciences, Inc.			
STREET: 9410 Key West Avenue			
CITY: Rockville			
STATE: Maryland			
COUNTRY: USA			
ZIP: 20850			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage			

```

1      COMPUTER:  HP Vectra 486/33
2      OPERATING SYSTEM:  MSDOS version 6.2
3      SOFTWARE:  ASCII Text
4
5      CURRENT APPLICATION DATA:
6      APPLICATION NUMBER:  US/09/765,271
7      FILING DATE:  22-Jan-2001
8      CLASSIFICATION:  <unknown>
9
10     PRIOR APPLICATION DATA:
11     APPLICATION NUMBER:  09/536,784
12     FILING DATE:  <unknown>
13     APPLICATION NUMBER:  08/961,083
14     FILING DATE:  OCT-30-1997
15
16     ATTORNEY/AGENT INFORMATION:
17     NAME:  Michelle S. Marks
18     REGISTRATION NUMBER:  41,971
19     REFERENCE/DOCKET NUMBER:  PB340P3
20     TELECOMMUNICATION INFORMATION:
21     TELEPHONE:  (301) 309-8504
22     TELEFAX:  (301) 309-8512
23
24     INFORMATION FOR SEQ ID NO: 1:
25     SEQUENCE CHARACTERISTICS:
26     LENGTH:  1999 base pairs
27     TYPE:  nucleic acid
28     STRANDEDNESS:  double
29     TOPOLOGY:  linear
30
31     SEQUENCE DESCRIPTION:  SEQ ID NO: 1:
32
33     US-09-765-271-1

```

Query Match	Similarity	100.0%	Score 1999:	DB 30:	Length 1999:
Best Local	Similarity	100.0%	Pred. No. 0:		
Matches 1999:	Conservative	0:	Mismatches	0:	Indels
					Gaps
QY	1	TAAATCTACGCAATATAAAATCAACTCATGTCGCTGACTTGGTCTTGATGACGCCGCGTCAA	60		
Db	1	TAAATCTACGCAATATAAAATCAACTCATGTCGCTGACTTGGTCTTGATGACGCCGCGTCAA	60		
QY	61	TGCCCAACCTAATGATATTTCCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACA	120		
Db	61	TGCCCAACCTAATGATATTTCCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACA	120		
QY	121	TGCGTCTTCGACACAGAGGGGATGATACATCCGATTCCTGGGAGGCTTCTTGCGCAA	180		
Db	121	TGCGTCTTCGACACAGAGGGGATGATACATCCGATTCCTGGGAGGCTTCTTGCGCAA	180		
QY	181	TCTGCAAGCAATTCCTCTCCAGGTGATCACTCTCACCCACAGTTGATTAACTTGAC	240		
Db	181	TCTGCAAGCAATTCCTCTCCAGGTGATCACTCTCACCCACAGTTGATTAACTTGAC	240		
QY	241	TTACTTTTCACTTGACACTTCGACACAGCTATTTCTGTAAAGGCTCGAAGAGCTTGGT	300		
Db	241	TTACTTTTCACTTGACACTTCGACACAGCTATTTCTGTAAAGGCTCGAAGAGCTTGGT	300		
QY	301	AGCGATTCAGTTAGAACAAAAAGACACCAAGCAAGAAATCTTGACCTACTATATATAATA	360		
Db	301	AGCGATTCAGTTAGAACAAAAAGACACCAAGCAAGAAATCTTGACCTACTATATATAATA	360		
QY	361	GCTCTACATGTCTAATGGGAACTATGAATGACAGACAGCTCAAAACTACTATATGTTAA	420		
Db	361	GCTCTACATGTCTAATGGGAACTATGAATGACAGACAGCTCAAAACTACTATATGTTAA	420		
QY	421	AGACCTCATATAATTTAAAGTTTACCTACAGTACCTTTGGTGGGTGGAAATGCGTCAGGAC	480		
Db	421	AGACCTCATATAATTTAAAGTTTACCTACAGTACCTTTGGTGGGTGGAAATGCGTCAGGAC	480		
QY	481	AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACGCGCGAAACTTGCTT	540		
Db	481	AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACGCGCGAAACTTGCTT	540		
QY	541	ATCTGAATGAAAAATCAAGGCTAATCTGTCGTGAACAGTTGAGAAAGCAGTGAATAC	600		
Db	541	ATCTGAATGAAAAATCAAGGCTAATCTGTCGTGAACAGTTGAGAAAGCAGTGAATAC	600		

QY	601	ACCAATTAAGTGAAGTCTCAAAATGACAACTAATTTACCGCTTACATGGA	660
Db	601	ACCAATTAAGTGAAGTCTCAAAATGACAACTAATTTACCGCTTACATGGA	660
QY	661	TAAATTACCTCAAGGAAGTCATCAATCAAGTTGGAAGAAAGAGGCTATTAACCTACTCAC	720
Db	661	TAAATTACCTCAAGGAAGTCATCAATCAAGTTGGAAGAAAGAGGCTATTAACCTACTCAC	720
QY	721	AACATGGATGGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
Db	721	AACATGGATGGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
QY	781	TTTCAATACAGACGAATACGTTGCTATATCAGACGATGTAATTTGCAAGTCGCTTACACAT	840
Db	781	TTTCAATACAGACGAATACGTTGCTATATCAGACGATGTAATTTGCAAGTCGCTTACACAT	840
QY	841	TGTTGATGTTTCTAACGGTAAAGTCATTGCGCCAGCTAGAGGACGCCATCAAGTCAAGTA	900
Db	841	TGTTGATGTTTCTAACGGTAAAGTCATTGCGCCAGCTAGAGGACGCCATCAAGTCAAGTA	900
QY	901	TGTTTCTTGGGAATTAACCAAGCAGTAGAACAACACGCGACTGGGATCACTAGTAA	960
Db	901	TGTTTCTTGGGAATTAACCAAGCAGTAGAACAACACGCGACTGGGATCACTAGTAA	960
QY	961	ACCGATACAGACGATATGCTGCTGGAGTAGAGGGTGCTACGATTAACCTGCTACAT	1020
Db	961	ACCGATACAGACGATATGCTGCTGGAGTAGAGGGTGCTACGATTAACCTGCTACAT	1020
QY	1021	CGTTACAGATGAGGCCCTATTAACCTACCCCTGGAGCAATTAACCTGTTTAACTGGGATAG	1080
Db	1021	CGTTACAGATGAGGCCCTATTAACCTACCCCTGGAGCAATTAACCTGTTTAACTGGGATAG	1080
QY	1081	GGGCTACTTTGGCAATCACTGCTTGCATATACGCTCTGCACAACTGCGAAACGTCCAGC	1140
Db	1081	GGGCTACTTTGGCAATCACTGCTTGCATATACGCTCTGCACAACTGCGAAACGTCCAGC	1140
QY	1141	CGTGAAGACTTAACAAGGTGGAGCTCAACCGCGCAGAGACTTTCGTAATGGTCTAGG	1200
Db	1141	CGTGAAGACTTAACAAGGTGGAGCTCAACCGCGCAGAGACTTTCGTAATGGTCTAGG	1200
QY	1201	AATGCACTACCCAAAGTATTCACACTCACTCAATAGCCATTTCAAGTAACACACCGAATCGA	1260
Db	1201	AATGCACTACCCAAAGTATTCACACTCACTCAATAGCCATTTCAAGTAACACACCGAATCGA	1260
QY	1261	CAAAAAATATGAGAGCAAGTAGTGAAGAAAGATGCGTCTGCTTACGCTTGGCAAAATGG	1320
Db	1261	CAAAAAATATGAGAGCAAGTAGTGAAGAAAGATGCGTCTGCTTACGCTTGGCAAAATGG	1320
QY	1321	TGGAACCTACTATAAACCAATGTATATCCATAAAGTCGCTTAAATGATGGAGATGAAAA	1380
Db	1321	TGGAACCTACTATAAACCAATGTATATCCATAAAGTCGCTTAAATGATGGAGATGAAAA	1380
QY	1381	AGATTTCTTAATGTCGGAACCTGTCCTATGGAAGAAAGACAGCGCTTATGATGACCGA	1440
Db	1381	AGATTTCTTAATGTCGGAACCTGTCCTATGGAAGAAAGACAGCGCTTATGATGACCGA	1440
QY	1441	CATGATGAAAAACAGTCTTGCACTTATGAGAACTGGAGAAATGCGTATCTTGGCTGCC	1500
Db	1441	CATGATGAAAAACAGTCTTGCACTTATGAGAACTGGAGAAATGCGTATCTTGGCTGCC	1500
QY	1501	TCAGAGCTGGTAAACAGGAACCTCTAATATACAGACGAGGAATTTGAAAAACACATCAA	1560
Db	1501	TCAGAGCTGGTAAACAGGAACCTCTAATATACAGACGAGGAATTTGAAAAACACATCAA	1560
QY	1561	GACCTTCATATTTGTAGACACTGATTAACCTATTTGGTGGCTTACGCGTAAATATTCAT	1620
Db	1561	GACCTTCATATTTGTAGACACTGATTAACCTATTTGGTGGCTTACGCGTAAATATTCAT	1620
QY	1621	GGCTGTATGAGACAGGCTATTTCAACCGTCTGACACACACTGTAGGCATATGCGCTTAAGGT	1680
Db	1621	GGCTGTATGAGACAGGCTATTTCAACCGTCTGACACACACTGTAGGCATATGCGCTTAAGGT	1680
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACTTAACCTGTCTGAGAGAGCAATCCAGANA	1740

Db	1681	CGCTGCCAAGTTTACCGCTCTATGATGACCTTACCTGTCTGAAAGGAGCAATCCAGAAGA	1740
Qy	1741	TTGGAATATATACCGAGAGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGCTGCTCG	1800
Db	1741	TTGGAAATATACCGAGAGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAAAGGTGCTCG	1800
Qy	1801	TTTCAGCTGGAACTCACCTCTCCACACACACCCCATCACTACGTGAAGTTCACGCTCATC	1860
Db	1801	TTTCAGCTGGAACTCACCTCTCCACACACACCCCATCACTACGTGAAGTTCACGCTCATC	1860
Qy	1861	ATCAGATAGTTCAACTTCACAGCTGTAGCTCAACCACTCCACACAAATATATAGTAGCAC	1920
Db	1861	ATCAGATAGTTCAACTTCACAGCTGTAGCTCAACCACTCCACACAAATATATAGTAGCAC	1920
Qy	1921	TACCAATCTCTAACATATATACGCAACCAATTCACACCCCTGTATCAACAAAAATTCAGAA	1980
Db	1921	TACCAATCTCTAACATATATACGCAACCAATTCACACCCCTGTATCAACAAAAATTCAGAA	1980
Qy	1981	TTCTCACCAGCAGCAACCA	1999
Db	1981	TTCTCACCAGCAGCAACCA	1999

```

RESULT 3
US-09-765-272-1
Sequence 1, Application US/09765272
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1999 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-765-272-1
Query Match 100.0%; Score 1999; DB 30; Length 1999;
Best Local Similarity 100.0%; Prid. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TAAATCTACGACATATAAATCAACATCTATTGCTACTTGGTTCTGAGCGCGCTCA 60
|||||

```


Db	1	TAATATCTACGACATATAAAATATCACTCATCTGCTGACTTGGGTTCTGAAACGCCGGCTCA	60
OY	61	TGCCCAAGCTATGATATATCTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACCA	120
Db	61	TGCCCAAGCTATGATATATCTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACCA	120
OY	121	TGCGTCTCTGACACACAGGGGGGATGTGATACATCCGATATCTGGGAGCTTCTTCCGCA	180
Db	121	TGCGTCTCTGACACACAGGGGGGATGTGATACATCCGATATCTGGGAGCTTCTTCCGCA	180
OY	181	TCTGCAAGCAATTCCTCCAGAGTGGATCACTCCACCAACAGTTGATTAAGTTATC	240
Db	181	TCTGCAAGCAATTCCTCCAGAGTGGATCACTCCACCAACAGTTGATTAAGTTATC	240
OY	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTTCTGTAAGGCTCAGGAAGCTTGGTT	300
Db	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTTCTGTAAGGCTCAGGAAGCTTGGTT	300
OY	301	AGCGATTCAGTTAGACAAAAGCAACCAAGCAGAATACTTTGACCTACTATATAATA	360
Db	301	AGCGATTCAGTTAGACAAAAGCAACCAAGCAGAATACTTTGACCTACTATATAATA	360
OY	361	GGGTACATGTCTAATGAGCAATGTGGAAATGCGACAGAGAGCTCAAAACTACTATGTA	420
Db	361	GGGTACATGTCTAATGAGCAATGTGGAAATGCGACAGAGAGCTCAAAACTACTATGTA	420
OY	421	AGACCTCAATTAATTAAGTTTACCTCAGTAGTGCCTTGCTGGTAATGCTCAGCAGCC	480
Db	421	AGACCTCAATTAATTAAGTTTACCTCAGTAGTGCCTTGCTGGTAATGCTCAGCAGCC	480
OY	481	AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCGCCGAACCTGGCTT	540
Db	481	AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCGCCGAACCTGGCTT	540
OY	541	ATTCGAATGAAAAATCAAGGCTACATCTGCTGTAAGCTGTGAAGAAAGCAAGCTAATAC	600
Db	541	ATTCGAATGAAAAATCAAGGCTACATCTGCTGTAAGCTGTGAAGAAAGCAAGCTAATAC	600
OY	601	ACCAATTAAGTATGACTATCAAAAGTCTCAAAATGAGCAAGTAATTAACCTGCTATACATGA	660
Db	601	ACCAATTAAGTATGACTATCAAAAGTCTCAAAATGAGCAAGTAATTAACCTGCTATACATGA	660
OY	661	TAAATTAACCTCAAGAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTATTAACCTACTAC	720
Db	661	TAAATTAACCTCAAGAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTATTAACCTACTAC	720
OY	721	AACGTGATGATATCTTCAACAAATGTGAACCAAGAGTCAAAAACATCTGGGATAT	780
Db	721	AACGTGATGATATCTTCAACAAATGTGAACCAAGAGTCAAAAACATCTGGGATAT	780
OY	781	TTTAATATACAGACAATACCTTGCTATCCAGACGATGAATTTGCAAGTGCCTTCAACAT	840
Db	781	TTTAATATACAGACAATACCTTGCTATCCAGACGATGAATTTGCAAGTGCCTTCAACAT	840
OY	841	TGTTGATGTCTTCAAGGTAAGTCAATGTGCCAGCTAGGAGCAGCCATCACTCAAGTAA	900
Db	841	TGTTGATGTCTTCAAGGTAAGTCAATGTGCCAGCTAGGAGCAGCCATCACTCAAGTAA	900
OY	901	TGTTGATGTCTTCAAGGTAAGTCAATGTGCCAGCTAGGAGCAGCCATCACTCAAGTAA	960
Db	901	TGTTGATGTCTTCAAGGTAAGTCAATGTGCCAGCTAGGAGCAGCCATCACTCAAGTAA	960
OY	961	ACCGATCAACAGATATGCTCTGCTGCTGGAGTACGCTGTCTACGANTCAACTGCTACTAT	1020
Db	961	ACCGATCAACAGATATGCTCTGCTGCTGGAGTACGCTGTCTACGANTCAACTGCTACTAT	1020
OY	1021	CGTTACAGATGAGCCCTATTACTACCTCGGACAAATTACTCTGTTTAACTGGGATAG	1080
Db	1021	CGTTACAGATGAGCCCTATTACTACCTCGGACAAATTACTCTGTTTAACTGGGATAG	1080
OY	1081	GGGCTACTTTGGCAACATCACTTGCATACGCCCTGCAACAAATGGCGAAAGTCCAGC	1140
Db	1081	GGGCTACTTTGGCAACATCACTTGCATACGCCCTGCAACAAATGGCGAAAGTCCAGC	1140

Qy	1141	CGTGGAAACCTCTAAACAAGCTGGACATCCAAACGGGCCAACAACCTTTCTTAATATGGTCTAGG	1200
Db	1141	CGTGGAAACCTCTAAACAAGCTGGACATCCAAACGGGCCAACAACCTTTCTTAATATGGTCTAGG	1200
Qy	1201	AATGCACTAACCCAGATATCTACTCTCAATATGCCATTTGCAAGTATACCAACGGAATCAGA	1260
Db	1201	AATGCACTAACCCAGATATCTACTCTCAATATGCCATTTGCAAGTATACCAACGGAATCAGA	1260
Qy	1261	CAAAAAATATGAGACCAAGTACTGAAAAAGATGGCTGCTCTTAAAGCTGGCTTTGCCAAATGG	1320
Db	1261	CAAAAAATATGAGACCAAGTACTGAAAAAGATGGCTGCTCTTAAAGCTGGCTTTGCCAAATGG	1320
Qy	1321	TGGAACTTACTATTAACCAATATCTATATCCATAAAGTGTCTTTAAGTATGGAGTGAATA	1380
Db	1321	TGGAACTTACTATTAACCAATATCTATATCCATAAAGTGTCTTTAAGTATGGAGTGAATA	1380
Qy	1381	AGAGTTCTCTAATATGCGCAATCTGGCCATGGAAGAAACGACAGCCATATATGATGACCGA	1440
Db	1381	AGAGTTCTCTAATATGCGCAATCTGGCCATGGAAGAAACGACAGCCATATATGATGACCGA	1440
Qy	1441	CATGATGAAAAACAGTCTTGACTTATGGAACCTGACGAATGCTATCTTGCTTGGCTCCC	1500
Db	1441	CATGATGAAAAACAGTCTTGACTTATGGAACCTGACGAATGCTATCTTGCTTGGCTCCC	1500
Qy	1501	TCAGGCTGTGTAACCAAGCAACCTTACTATCTATACGACAGAGAGAAATGGAACCAACATCA	1560
Db	1501	TCAGGCTGTGTAACCAAGCAACCTTACTATCTATACGACAGAGAGAAATGGAACCAACATCA	1560
Qy	1561	GACCTCTCAATTTGTAGACCTGATGAAATATTTGCTGGCTATACGGGTAAATATTTCAAT	1620
Db	1561	GACCTCTCAATTTGTAGACCTGATGAAATATTTGCTGGCTATACGGGTAAATATTTCAAT	1620
Qy	1621	GAGCTGATGGAAGGCTATTTCTTAACGCTGTGACACCACTTGTAGCCAAATGGCTTAACGGT	1680
Db	1621	GAGCTGATGGAAGGCTATTTCTTAACGCTGTGACACCACTTGTAGCCAAATGGCTTAACGGT	1680
Qy	1681	CGCTGCCAAAGTTTAAACGCTCTATGATGAGACCTACCTGCTGGAAGGAAGCAATCCAGACGA	1740
Db	1681	CGCTGCCAAAGTTTAAACGCTCTATGATGAGACCTACCTGCTGGAAGGAAGCAATCCAGACGA	1740
Qy	1741	TTTGAATATACCAAGAGGGGCTCTACAGAAATGGAATTTGCTATTTAAAAATGGTGTCTCG	1800
Db	1741	TTTGAATATACCAAGAGGGGCTCTACAGAAATGGAATTTGCTATTTAAAAATGGTGTCTCG	1800
Qy	1801	TTTCTAGTGGAACTACCTGCTCCACAAACACCCCATATCACTGAAAGTTCAAGCTCATC	1860
Db	1801	TTTCTAGTGGAACTACCTGCTCCACAAACACCCCATATCACTGAAAGTTCAAGCTCATC	1860
Qy	1861	ATTCAGATGTGTTCAACTTACAGCTTACTCTCAACCAATCCCAAGCAACAATATATGTATCGAC	1920
Db	1861	ATTCAGATGTGTTCAACTTACAGCTTACTCTCAACCAATCCCAAGCAACAATATATGTATCGAC	1920
Qy	1921	TACCAATCTCTAAACAATAATACGAACAATCAAAATATCAACCCCTGATCAACAAATCAGAA	1980
Db	1921	TACCAATCTCTAAACAATAATACGAACAATCAAAATATCAACCCCTGATCAACAAATCAGAA	1980
Qy	1981	TTCTTCAACGACACACCA 1999	
Db	1981	TTCTTCAACGACACACCA 1999	

RESULT 4

US-60-029-960-81/c

Sequence 81, Application US/60029960

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 1649

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

```

RESULT      4
US-60-029-960-81/c
: Sequence 81, Application US/60029960
:
: GENERAL INFORMATION:
:
: APPLICANT: Charles Kunsch
:
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
:
: NUMBER OF SEQUENCES: 1643
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Human Genome Sciences, Inc.
:
: STREET: 9410 Key West Avenue
:
: CITY: Rockville
:

```

STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/60/029,960
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: PB340PP
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 81:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10383 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-60-029-960-81

Query Match 100.0%; Score 1999; DB 41; Length 10383;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAAATCTACGCAATTAATAATCAATCTATTGCTGCTTGGTTCTGAAGCCGCGTCAA 60
 Db 9763 TAAATCTACGCAATTAATAATCAATCTATTGCTGCTTGGTTCTGAAGCCGCGTCAA 9704
 OY 61 TCCCAAGTAATGATATTTCCACAGATTGGTTAAGCAATCGTTTCATGGAAGACCA 120
 Db 9703 TCCCAAGTAATGATATTTCCACAGATTGGTTAAGCAATCGTTTCATGGAAGACCA 9644
 OY 121 TCGCTTCTCGACACAGAGGGGATGATACATCCGATCTCGGAGCTTCTTCGCGCA 180
 Db 9643 TCGCTTCTCGACACAGAGGGGATGATACATCCGATCTCGGAGCTTCTTCGCGCA 9584
 OY 181 TCTGCAAGCAATTCCTTCGCAAGGTGATCACTCTACCCCAACGTTGATTAACTGAC 240
 Db 9583 TCTGCAAGCAATTCCTTCGCAAGGTGATCACTCTACCCCAACGTTGATTAACTGAC 9524
 OY 241 TTACTTTCACTTCGACTCCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGTT 300
 Db 9523 TTACTTTCACTTCGACTCCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGTT 9464
 OY 301 AGCGATTAGTAGACAAAAAGCAACAGCAAGAAATCTGACCTATATATAATA 360
 Db 9463 AGCGATTAGTAGACAAAAAGCAACAGCAAGAAATCTGACCTATATATAATA 9404
 OY 361 GGTCTACATGCTTAATGGAAGTATGAATGAGACAGAGCTCAAAAATCTATGTAA 420
 Db 9403 GGTCTACATGCTTAATGGAAGTATGAATGAGACAGAGCTCAAAAATCTATGTAA 9344
 OY 421 AGACTCAATATTAATGCTTACCTCAGTATGCTTGTGCTGGAATGCCCTCAGGAC 480
 Db 9343 AGACTCAATATTAATGCTTACCTCAGTATGCTTGTGCTGGAATGCCCTCAGGAC 9284
 OY 481 AAACCAATATGACCCCTATTTCATCATCAGAAAGCAGCCCAAGACCCGAAATTTGGT 540
 Db 9283 AAACCAATATGACCCCTATTTCATCATCAGAAAGCAGCCCAAGACCCGAAATTTGGT 9224
 OY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAAAGCAGTCAATAC 600

Db 9223 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAAAGCAGTCAATAC 9164
 OY 601 ACCAATTAATGATGAGCTACCAAGTCTCAATACAGCAAGTAAATTAATCCCTGTTACATGGA 660
 Db 9163 ACCAATTAATGATGAGCTACCAAGTCTCAATACAGCAAGTAAATTAATCCCTGTTACATGGA 9104
 OY 661 TAATTAATCTCAAGGAAGTATCATCAATCAAGTGAAGGAAGCAAGGCTATTAACCTACTAC 720
 Db 9103 TAATTAATCTCAAGGAAGTATCATCAATCAAGTGAAGGAAGCAAGGCTATTAACCTACTAC 9044
 OY 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAATCTGTTGGATAT 780
 Db 9043 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAATCTGTTGGATAT 8984
 OY 781 TTACATACAGACGAATAGTTCCTATCCAGACATGATGATGCAAGTCCCTTACCAT 840
 Db 8983 TTACATACAGACGAATAGTTCCTATCCAGACATGATGATGCAAGTCCCTTACCAT 8924
 OY 841 TGTGATGTTTCTTAACGGTAAAGTCAATGCGCAGCTAGAGACGCCATCACTCAAGTAA 900
 Db 8923 TGTGATGTTTCTTAACGGTAAAGTCAATGCGCAGCTAGAGACGCCATCACTCAAGTAA 8864
 OY 901 TGTTCCTTCCGAATTAACCAAGCAGTAGAAACCAACCCGACGCTGGGATCACTATGAA 960
 Db 8863 TGTTCCTTCCGAATTAACCAAGCAGTAGAAACCAACCCGACGCTGGGATCACTATGAA 8804
 OY 961 ACCGATACAGACTATAGCTCCTGCTGGAGTAGGTGCTCTAGATCACTCACTACTAT 1020
 Db 8803 ACCGATACAGACTATAGCTCCTGCTGGAGTAGGTGCTCTAGATCACTCACTACTAT 8744
 OY 1021 CGTTCAGATGAGCCCTATTAATCACTACCTGGGACAAATACCTCTGTTAATCTGGGATAG 1080
 Db 8743 CGTTCAGATGAGCCCTATTAATCACTACCTGGGACAAATACCTCTGTTAATCTGGGATAG 8684
 OY 1081 GGGCTACTTTGGCAACATCACTTGCATATACGCCCTCGCAACAATGGCCAAAGCTCCAGC 1140
 Db 8683 GGGCTACTTTGGCAACATCACTTGCATATACGCCCTCGCAACAATGGCCAAAGCTCCAGC 8624
 OY 1141 CGTGAACACTTAACAAGGTGGAGCTCAACCGGCCCAAGACTTTCCTAATAGTGTAGG 1200
 Db 8623 CGTGAACACTTAACAAGGTGGAGCTCAACCGGCCCAAGACTTTCCTAATAGTGTAGG 8564
 OY 1201 AATGACTACCCAGTATTAATCACTCAATGCCATTTCAATTAACACACCGAATCAGA 1260
 Db 8563 AATGACTACCCAGTATTAATCACTCAATGCCATTTCAATTAACACACCGAATCAGA 8504
 OY 1261 CAAAAAATATGAGCAAGTAGAGAAAGATGCTGCTTACGCTGCCCTTGGCAATGG 1320
 Db 8503 CAAAAAATATGAGCAAGTAGAGAAAGATGCTGCTTACGCTGCCCTTGGCAATGG 8444
 OY 1321 TGAAGCTTACTTAACCAATGATATATGCAATGAAGCTGCTTACGCTGCCCTTGGCAATGG 1380
 Db 8443 TGAAGCTTACTTAACCAATGATATATGCAATGAAGCTGCTTACGCTGCCCTTGGCAATGG 8384
 OY 1381 AGAGTTCTTAATGTCGGAAGTCTGCAATGAAGAAAGACAGGCTATATATGACCGA 1440
 Db 8383 AGAGTTCTTAATGTCGGAAGTCTGCAATGAAGAAAGACAGGCTATATATGACCGA 8324
 OY 1441 CATGATGAAGACAGTCTTATGATGAATGAGACGAATGCTATCTTGGCTGCC 1500
 Db 8323 CATGATGAAGACAGTCTTATGATGAATGAGACGAATGCTATCTTGGCTGCC 8264
 OY 1501 TCAGGCTGTTAAACAGGACCTTAACTATACAGCGGAATTTGAAACCAATCA 1560
 Db 8263 TCAGGCTGTTAAACAGGACCTTAACTATACAGCGGAATTTGAAACCAATCA 8204
 OY 1561 GACCTCTCAATTTTGAACGCTGATGAATTTTCTGCTATTAACGCTTAATATTCAT 1620
 Db 8203 GACCTCTCAATTTTGAACGCTGATGAATTTTCTGCTATTAACGCTTAATATTCAT 8144
 OY 1621 GGCTGATGAGAGGCTATTTCAACGCTGTGACACCACTTTGAGCAATGGCTTACGCT 1680
 Db 8143 GGCTGATGAGAGGCTATTTCAACGCTGTGACACCACTTTGAGCAATGGCTTACGCT 8084

```

OY 1681 CGCGCCAAATTTTACCGCTGTATGATGAGACTACTGTCTGGAAGGAAGCAATCCAGAA 1740
Db 8083 CGCTGCCAAAGTTTACCGCTGTATGATGAGACTACTGTCTGGAAGGAAGCAATCCAGAA 8024
OY 1741 TTGGATATATACAGAGGGGCTCTACAGAAATGAGAAATTGATTTTAAAAATGGTCTCG 1800
Db 8023 TTGGATATATACAGAGGGGCTCTACAGAAATGAGAAATTGATTTTAAAAATGGTCTCG 7964
OY 1801 TTTCACGTGGAACTCACCCTGCTCCACAACAACCCCAATCACTGAAAGTTCAAGCTCATC 1860
Db 7963 TTTCACGTGGAACTCACCCTGCTCCACAACAACCCCAATCACTGAAAGTTCAAGCTCATC 7904
OY 1861 ATCAGATAGTTCAACTCTCACAGTCTACCTCAACACCTCAAGACAATAATATAGTACGAC 1920
Db 7903 ATCAGATAGTTCAACTCTCACAGTCTACCTCAACACCTCAAGACAATAATATAGTACGAC 7844
OY 1921 TACCAATCTTACAAATATATAGCAACAAATCAATCAACCCCTGATCAACAAATCAGAA 1980
Db 7843 TACCAATCTTACAAATATATAGCAACAAATCAATCAACCCCTGATCAACAAATCAGAA 7784
OY 1981 TCTCTAACCCAGCACACCA 1999
Db 7783 TCTCTAACCCAGCACACCA 7765

RESULT 5
US-08-961-527-145/C
: Sequence 145, Application US/08961527
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 391
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,527
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEO ID NO: 145:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 10711 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-961-527-145

Query Match 100.0%; Score 1999; DB 13; Length 10711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TAAAAATCTACGACATATAAAAATCAACTCATTTGCTGACTTGCGTTCTGAAACGCCGCGCTCA 60

```

D	b	9767	TA	AATCTACGACAACTAAAAATCAGCTCATCTGCTGACTTGGTTCTGAAAGCCGGCTCA	9708
O	y	61	TC	CCCAAGCTAATGATATTCOCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACCA	120
D	b	9707	TC	CCCAAGCTAATGATATTCOCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACCA	9648
O	y	121	TC	GGTTCCTGCACACAGGGGGGATGTATACATCCGATTCCTGGGGGCTTCTTCGGCA	180
D	b	9647	TC	GGTTCCTGCACACAGGGGGGATGTATACATCCGATTCCTGGGGGCTTCTTCGGCA	9588
O	y	181	TC	TGCAAAAGCAATTCCTCCAGGTGATCACTCTCACCAACACATTTGATTAAGTTGAC	240
D	b	9587	TC	TGCAAAAGCAATTCCTCCAGGTGATCACTCTCACCAACACATTTGATTAAGTTGAC	9528
O	y	241	TT	ACTTTTCAACTTCGACCTCCGACCAAGACATATTTTCGTAAAGGCTCAGAGCTTTGGT	300
D	b	9527	TT	ACTTTTCAACTTCGACCTCCGACCAAGACATATTTTCGTAAAGGCTCAGAGCTTTGGT	9468
O	y	301	AG	CGATTCAGTTAGACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATTAATTA	360
D	b	9467	AG	CGATTCAGTTAGACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATTAATTA	9408
O	y	361	GG	TCTACATGCTCTATGAGGAATATGGAATGCAGACAGACGCTCAAACTACTATATGTTA	420
D	b	9407	GG	TCTACATGCTCTATGAGGAATATGGAATGCAGACAGACGCTCAAACTACTATATGTTA	9348
O	y	421	AG	ACCTCAATTAATTTAGTTTACCTCAAGTTAGCTTGCGGCGGGAATCCCGACGACCC	480
D	b	9347	AG	ACCTCAATTAATTTAGTTTACCTCAAGTTAGCTTGCGGCGGGAATCCCGACGACCC	9288
O	y	481	AA	ACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCGCCGAAACTTGGTCTT	540
D	b	9287	AA	ACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCGCCGAAACTTGGTCTT	9228
O	y	541	AT	CTGAAATGAAAAATTAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600
D	b	9227	AT	CTGAAATGAAAAATTAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	9168
O	y	601	AC	CAATTAACGATGACTAGACAAGGTCTCAAAATGACAGAAATTAATTAACCGCTTACATGA	660
D	b	9167	AC	CAATTAACGATGACTAGACAAGGTCTCAAAATGACAGAAATTAATTAACCGCTTACATGA	9108
O	y	661	TA	ATTAACCTCAAGAGAGTCAATCAATCAAGTTGAAGAAAGCAAGCGCTATTAACCTACTAC	720
D	b	9107	TA	ATTAACCTCAAGAGAGTCAATCAATCAAGTTGAAGAAAGCAAGCGCTATTAACCTACTAC	9048
O	y	721	AA	CTGGGATGGATGCTCTACACAATGATGAGACCAAGAGTCAAAAACATCTGTGGGATAT	780
D	b	9047	AA	CTGGGATGGATGCTCTACACAATGATGAGACCAAGAGTCAAAAACATCTGTGGGATAT	8988
O	y	781	TT	ACAATACAGACGAATAGCTTGCTCATCCAGACGATGAATTTGCAAGTCGTTTACCAT	840
D	b	8987	TT	ACAATACAGACGAATAGCTTGCTCATCCATCCAGACGATGAATTTGCAAGTCGTTTACCAT	8928
O	y	841	TT	GTGATGTTTCTAAGCGTAAAGTCAATTCGCTCCAGTAGAGACCAAGCAAGTCAAGTAA	900
D	b	8927	TT	GTGATGTTTCTAAGCGTAAAGTCAATTCGCTCCAGTAGAGACCAAGCAAGTCAAGTAA	8868
O	y	901	TC	TTTCCTTCTGGAAATTAACCAACGACATAGAAACCAACCCGACTGGGATCAACTATGAA	960
D	b	8867	TC	TTTCCTTCTGGAAATTAACCAACGACATAGAAACCAACCCGACTGGGATCAACTATGAA	8808
O	y	961	AC	CGATACAGACTATGCTCTGCTCGCTTGGAGTACGGGTCTACAGATTCACACTGCTACTAT	1020
D	b	8807	AC	CGATACAGACTATGCTCTGCTCGCTTGGAGTACGGGTCTACAGATTCACACTGCTACTAT	8748
O	y	1021	CG	TTCACAGATGAGGCCCTATATACATACCTCGGAGCAAAATTAATCTCTGTTTAACTGGAGTAG	1080
D	b	8747	CG	TTCACAGATGAGGCCCTATATACATACCTCGGAGCAAAATTAATCTCTGTTTAACTGGAGTAG	8688
O	y	1081	GG	GTACTTTTGGACAATCACTTGTGAATACGCCCTGCAACAAATGCGAAAGCTGCCAGC	1140

D	8687	GGGCTACTTGGCAACATCATCCCTTGCATTAACGCCCTGGCAACATTCGGGAACAGCTCCAGC	8628
Q	1141	CGTGGAACTCTAACAAGGTGGACTCAACCGCGCCAGACCTTTCTCTAAATGGTCTAGG	1200
D	8627	CGTGGAACTCTAACAAGGTGGACTCAACCGCGCCAGACCTTTCTCTAAATGGTCTAGG	8568
Q	1201	AATCGACTCCCAAGTATTCACCTCAACCAATGGCATTTTCAAGTAACCAACCGCAATCAGA	1260
D	8567	AATCGACTCCCAAGTATTCACCTCAACCAATGGCATTTTCAAGTAACCAACCGCAATCAGA	8508
Q	1261	CAAAAAATATGAGCAAGTAGTGAAGAAGATGGCTGCTTACGCTGACCTTTGCCAATG	1320
D	8507	CAAAAAATATGAGCAAGTAGTGAAGAAGATGGCTGCTTACGCTGACCTTTGCCAATG	8448
Q	1321	TGGAACTTACTATTAACCAATGATATATCCATAAAGTGGCTTATGATGAGGAGTGA	1380
D	8447	TGGAACTTACTATTAACCAATGATATATCCATAAAGTGGCTTATGATGAGGAGTGA	8388
Q	1381	AACAGTTCTTATATGTCGGAACCTGTCCTATGAAAGGAACGACAGCCTATATGATGACCGA	1440
D	8387	AACAGTTCTTATATGTCGGAACCTGTCCTATGAAAGGAACGACAGCCTATATGATGACCGA	8328
Q	1441	CATGATGAAAACAGTCTTACCTATATGAACTGGACGAAAATGCCATTCCTGCTGGGCTCC	1500
D	8327	CATGATGAAAACAGTCTTACCTATATGAACTGGACGAAAATGCCATTCCTGCTGGGCTCC	8268
Q	1501	TCAGGCTGGTAAACAGAGAACCTCTACTATACAGACGAGGAATTTGAAAACACATCAA	1560
D	8267	TCAGGCTGGTAAACAGAGAACCTCTACTATACAGACGAGGAATTTGAAAACACATCAA	8208
Q	1561	GACCTCTCAATTTGTAGCACCTGATGAACTATTTGCTGGCTATAGCGTAATATTCAT	1620
D	8207	GACCTCTCAATTTGTAGCACCTGATGAACTATTTGCTGGCTATAGCGTAATATTCAT	8148
Q	1621	GGCTGATGAGAGGGTATTTCAACCGTCGACACACACTGTTGGGAATGGGCTTACGGT	1680
D	8147	GGCTGATGAGAGGGTATTTCAACCGTCGACACACACTGTTGGGAATGGGCTTACGGT	8088
Q	1681	CGCTGCCAAGTTTACCGGCTTATGATGACCTACCTGCTGTGAAGGAAGCAATCCAGAGA	1740
D	8087	CGCTGCCAAGTTTACCGGCTTATGATGACCTACCTGCTGTGAAGGAAGCAATCCAGAGA	8028
Q	1741	TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTCTCG	1800
D	8027	TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTCTCG	7968
Q	1801	TTTCTACGAGGAACTACCGCTCCACCAACACCCCATCACTCAAGTAAGTTCAAGCTCATC	1860
D	7967	TTTCTACGAGGAACTACCGCTCCACCAACACCCCATCACTCAAGTAAGTTCAAGCTCATC	7908
Q	1861	ATCAGATAGTTTCACTTCACAGTCTAGCTCAACACCACTCCACGACACAATTAATAGTAGCAG	1920
D	7907	ATCAGATAGTTTCACTTCACAGTCTAGCTCAACACCACTCCACGACACAATTAATAGTAGCAG	7848
Q	1921	TACCAATCTTAACAATATAGCGACAAATCAAAATACACCCCTGATCAACAAATACAGAA	1980
D	7847	TACCAATCTTAACAATATAGCGACAAATCAAAATACACCCCTGATCAACAAATACAGAA	7788
Q	1981	TCCTCAACGACGACAACCA	1999
D	7787	TCCTCAACGACGACAACCA	7769

```

RESULT      6
US-09-583-110-1312
: Sequence 1312, Application US/09583110
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al.
: TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
: TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
: FILE REFERENCE: P4TH00-07A
: CURRENT APPLICATION NUMBER: US/09/583,110
: CURRENT FILING DATE: 2000-05-26

```

```

1 PRIOR APPLICATION NUMBER: US 09/107,433
2 PRIOR FILING DATE: 1998-06-30
3 PRIOR APPLICATION NUMBER: US 60/085,131
4 PRIOR FILING DATE: 1998-05-12
5 PRIOR APPLICATION NUMBER: US 60/051,553
6 PRIOR FILING DATE: 1997-07-02
7 NUMBER OF SEQ ID NOS: 5322
8 SEQ ID NO 1312
9 LENGTH: 2160
10 TYPE: DNA
11 ORGANISM: Streptococcus pneumoniae
12 US-09-563-110-1312

```

	Query Match	Similarity	99.4%	Score	1987.8	DB	22	Length	2160	
	Best Local	Similarity	99.6%	Pred. No.	0					
	Matches	1992	Conservative	0	Mismatches	7	Indels	0	Gaps	0
QY	1	TAAATCTAGACACATAAAAAATCAACTCATTTGCTGACTTGGGTTCTGAGACGGCGCTCAA	60							
Db	159	taaatcttagacacataaataaatacaactcatctgctacttggtctctgtaagccgctca	218							
QY	61	TGCCCAACTATATATATTTCCCAACAGATTTCGTTAAGGCATTCGTTTCTATACGAAGCCA	120							
Db	219	tgcocaaagctaatgatattcccaagaatttggtaaaggcaatcgltctatcgaagacca	278							
QY	121	TGCGTTCTTCACACACACAGGGGATTTGATACCATCGATTCCGTGGGAGGCTTCTTCGCCAA	180							
Db	279	tcgctctctgcagccacaggggattgatataccatccgatctcttggaagctctcttcgcgaa	338							
QY	181	TCTGCAAGAACAATTCCTCCACAGGTGATCACTCACTCAACCAACAGTTGATTAACTTGAC	240							
Db	339	tctgaaagaacatctccctccaaggttggtctcaactctcactcaaacagttgattaagttgac	398							
QY	241	TTACTTTTCAACCTTCGATTTCCGACACACATTTCTGTGAAGCTCAGAGAAGCTTGTT	300							
Db	399	ttaactttccaactctcgactctccgaccagactattctctgtaagctcagaagcttggct	458							
QY	301	AGCGATTTCAGTTAGAACAAAAAAGCAACCCAAAGAAATCTTGACCTACTATATATAATAA	360							
Db	459	agcgattcagtttagaacacaaagaacccaagaagaatcttgactactactataaataa	518							
QY	361	GCTCTACACTGCTCTAATGGGAACTATGGAATGCACACACACACTCAAAACTCTATGGTAA	420							
Db	519	ggtctaacctgtctaattgggaactatlgaaatgacagacagctcaaaactactatggtaa	578							
QY	421	AGACCTCATATATTAACTTACCTCACTGACTTGGCTGGCTGGAAATGCTTCAGGACCC	480							
Db	579	agacctcaataatttaagtttaactcagttagctcttgcttggttggaatgctctcaggaacc	638							
QY	481	AAACCAATATACCCCTTATTACATCCAGAAAGCAACCCCAACCCGGGAACTTGCTT	540							
Db	639	aaacccaataatgacccctattcacatccagaagcagcccaaacccgcgaaacttggctct	698							
QY	541	ATTGGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGACAGTCAATAC	600							
Db	699	atctggaatgaaaataatcgaagctatactctctgcctgaaacagatatagaagaagatcaat	758							
QY	601	ACCAATTACTGATGACTTACAAAGCTCTCAAAATACGCAAGTATATTACCTGCTTACATGA	660							
Db	759	accaattactgattggtctacaagaatctcaaatcagcaagtaattaccctcttacaatgga	818							
QY	661	TTAATTACCTCAAGGAAGCATCATCAAGTTGAGGAAGAAACAGGCTATTAACCTATCCAC	720							
Db	819	taattaccctcaaggaagatcacaatcaagttgaaagaagaacaggtcataaccctaccac	878							
QY	721	AACGTGGATGATTTCTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT	780							
Db	879	aactggtgattgattctacacaaatgtagaccagaagatcctcaaaacatcttgtyggtat	938							
QY	781	TTACAAATACAGACCAATACGTTGGCTTATCCAGACGATGATTTGCAAGTGGCTTACCAT	840							
Db	939	ttaacaatacagaacgaataacgttgcctctccagacgaatgaaatgtaagctcttcaact	998							

Db 225 TGCCCAAGCTTAATGATATTTCCACAGATTTGGTTAAGGCAATCTGTTCTATCGAAGACCA 284
Oy 121 TCGCTTCTTCGACACAGGGGGATGATACCATCCGTATCCCTGGAGGCTTTCTGGCAA 180
Db 285 TCGCTTCTTCGACACAGGGGGATGATACCATCCGTATCCCTGGAGGCTTTCTGGCAA 344
Oy 181 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTCACCCCAACAGTTGATTAACTGAC 240
Db 345 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTCACCCCAACAGTTGATTAACTGAC 404
Oy 241 TTACTTTCACTTGGACTTCGACAGACATTTCTGTAAGGCTCAGAAAGCTTGGT 300
Db 405 TTACTTTCACTTGGACTTCGACAGACATTTCTGTAAGGCTCAGAAAGCTTGGT 464
Oy 301 AGCGATTAGTAGAACAACCAAGCAGAGAAATCTGACCTACTATATATATA 360
Db 465 AGCGATTAGTAGAACAACCAAGCAGAGAAATCTGACCTACTATATATATA 524
Oy 361 GGTCTACATGTTCTAATGGGAATCTATGGAATGACAGACAGCTCAAAACTACTATGTA 420
Db 525 GGTCTACATGTTCTAATGGGAATCTATGGAATGACAGACAGCTCAAAACTACTATGTA 584
Oy 421 AGACTCATATTTAATTAATTTACCTCAGTTAGCTTGGCTGGATGCTCAGGACAC 480
Db 585 AGACTCATATTTAATTAATTTACCTCAGTTAGCTTGGCTGGATGCTCAGGACAC 644
Oy 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCAGCCGAACTTGGCT 540
Db 645 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCAGCCGAACTTGGCT 704
Oy 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAGAGATGAGAAAGCAGTCAATAC 600
Db 705 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAGAGATGAGAAAGCAGTCAATAC 764
Oy 601 ACCATTTCTGATGATGACATCAAGATCAATCAGCAAGTATTTCCCTGCTTAATGTA 660
Db 765 ACCATTTCTGATGATGACATCAAGATCAATCAGCAAGTATTTCCCTGCTTAATGTA 824
Oy 661 TAATTAATCTCAAGAAAGTATCAATCAAGTTGAGAAAGAGAGCTTAATCTCTAC 720
Db 825 TAATTAATCTCAAGAAAGTATCAATCAAGTTGAGAAAGAGAGCTTAATCTCTAC 884
Oy 721 AACTGGATGATGTTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTTGGAATAT 780
Db 885 AACTGGATGATGTTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTTGGAATAT 944
Oy 781 TTACATATCAGACGAATACGTTGCTATCCAGAGATGAATTTGCAAGTCCCTTACCAT 840
Db 945 TTACATATCAGACGAATACGTTGCTATCCAGAGATGAATTTGCAAGTCCCTTACCAT 1004
Oy 841 TGTGATGTTCTTAACGGTAAAGTATGCTGAGCTAGAGAGAGCTCAAGTCAAGTAA 900
Db 1005 TGTGATGTTCTTAACGGTAAAGTATGCTGAGCTAGAGAGAGCTCAAGTCAAGTAA 1064
Oy 901 TGTGATGTTCTTAACGGTAAAGTATGCTGAGAGAGAGCTCAAGTCAAGTAA 960
Db 1065 TGTGATGTTCTTAACGGTAAAGTATGCTGAGAGAGAGCTCAAGTCAAGTAA 1124
Oy 961 ACCGATCAGACATATGCTGCTGCTTGGAGTACGTTGATCAATTTCAATGCTACTAT 1020
Db 1125 ACCGATCAGACATATGCTGCTGCTTGGAGTACGTTGATCAATTTCAATGCTACTAT 1184
Oy 1021 CGTTACAGATAGCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
Db 1185 CGTTACAGATAGCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1244
Oy 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCTTGCAACATGCGAAAGCTCCAGC 1140
Db 1245 GGGCTACTTTGGCAACATCACCTTGCAATACGCTTGCAACATGCGAAAGCTCCAGC 1304
Oy 1141 CGTGAACCTTAACAAAGGTGAGCTACAGCGCGCAAGACTTCTTAATTTGCTCTAG 1200
Db 1305 CGTGAACCTTAACAAAGGTGAGCTACAGCGCGCAAGACTTCTTAATTTGCTCTAG 1364

Oy 1201 AATGACTACCCAAATATTTCACTACTCAAAATGCCATTTCAAGTAAACAGACGAATCAGA 1260
Db 1365 AATGACTACCCAAATATTTCACTACTCAAAATGCCATTTCAAGTAAACAGACGAATCAGA 1424
Oy 1261 CAAAAAATATGAGCAGAGTATGTAAGAGATGCTGCTCTTACGCTGCTTGGCAATG 1320
Db 1425 CAAAAAATATGAGCAGAGTATGTAAGAGATGCTGCTCTTACGCTGCTTGGCAATG 1484
Oy 1321 TGGACTTACTATTAACCAATATATCATTAATGCTGCTTATGATGAGAGTAA 1380
Db 1485 TGGACTTACTATTAACCAATATATCATTAATGCTGCTTATGATGAGAGTAA 1544
Oy 1381 AGAGTTCTTAATGTCGGAACCTGTCATGAGAGAAAGCAGAGCTTATGATGAGC 1440
Db 1545 AGAGTTCTTAATGTCGGAACCTGTCATGAGAGAAAGCAGAGCTTATGATGAGC 1604
Oy 1441 CATGATGAAACAGTCTTACTATGATGAGAGTGAAGTATGCTTGGCTGCC 1500
Db 1605 CATGATGAAACAGTCTTACTATGATGAGAGTGAAGTATGCTTGGCTGCC 1664
Oy 1501 TCAGGCTGTTAAACAGGACCTTACTATACAGAGGAAATGAAACCATCA 1560
Db 1665 TCAGGCTGTTAAACAGGACCTTACTATACAGAGGAAATGAAACCATCA 1724
Oy 1561 GACCTCTCAATTTGAGACCTGATGATATTTGCTGCTATACGCTAAATATTCAT 1620
Db 1725 GACCTCTCAATTTGAGACCTGATGATATTTGCTGCTATACGCTAAATATTCAT 1784
Oy 1621 GGCTGTATGACAGGCTATTTCAACGCTGACACCACTTGTAGCAATGGCTTACGCT 1680
Db 1785 GGCTGTATGACAGGCTATTTCAACGCTGACACCACTTGTAGCAATGGCTTACGCT 1844
Oy 1681 GGCTGTATGACAGGCTATTTCAACGCTGACACCACTTGTAGCAATGGCTTACGCT 1740
Db 1845 GGCTGTATGACAGGCTATTTCAACGCTGACACCACTTGTAGCAATGGCTTACGCT 1904
Oy 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCATTTAAATTTGCTG 1800
Db 1905 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCATTTAAATTTGCTG 1964
Oy 1801 TTTCAGTGAACCTACCTGCTCCACAAACCCCATCAACCTGAAGTCAAGCTATC 1860
Db 1965 TTTCAGTGAACCTACCTGCTCCACAAACCCCATCAACCTGAAGTCAAGCTATC 2024
Oy 1861 ATCAGATGTTCACTTACAGTCTAGCTCAACCACTCAAGCAATTAATGAGAC 1920
Db 2025 ATCAGATGTTCACTTACAGTCTAGCTCAACCACTCAAGCAATTAATGAGAC 2084
Oy 1921 TACCAATCTTAACATATATAGCAAAATCAATCAACCCCTGATCAACAAATCAGAA 1980
Db 2085 TACCAATCTTAACATATATAGCAAAATCAATCAACCCCTGATCAACAAATCAGAA 2144
Oy 1981 TCTCTAACCCAGCACACA 1999
Db 2145 TCTCTAACCCAGCACACA 2163

RESULT 8
PCT-US02-03987-9325
Sequence 9325, Application PC/TUS0203987
GENERAL INFORMATION:
APPLICANT: Elittra Pharmaceuticals, Inc.
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
FILE REFERENCE: ELITRA-028VPC
CURRENT APPLICATION NUMBER: PCT/US02/03987
PRIOR FILING DATE: 2002-02-02
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9325

Db 2139 tccttaaccagcacacca 2157

RESULT 9

US-09-815-242-9325

Sequence 9325, Application US/09815242

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21*

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9325

LENGTH: 2160

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(2160)

US-09-815-242-9325

Query Match 99.18; Score 1981.4; DB 31; Length 2160;

Best Local Similarity 99.48; Pred. No. 0;

Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATTCAGACAAATTAATCAATGCTGACTGGGTTCTGAAGCGCGGTCA 60

Db 159 taatatcagcaataaataacatcaatcgtgacttggtcttgagacgcgcgtcaa 218

QY 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGCAATCGTTTCTATCAAGACA 120

Db 219 tgcacaagcataatgatattccacagatttggttaaggcaatcglttctacagaagaca 278

QY 121 TCGCTTCTTGACACAGAGGGGATTTGATACATCCGTAATCTGGAGAGCTTTTGGCGAA 180

Db 279 tcgcttctcgcacacaggggagattgataccatccgatactctggaagcttcttcgcgcaa 338

QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTTCACCCCAACAGTTGATTAGTTGAC 240

Db 339 tctgcaagaacattccctccaaggltgaltcagcttcaactcaacagltgalttaagtac 398

QY 241 TTATTTTCAACTTCGACTTCGACAGACTTCTGTAAGGCTGAGAGAGCTTGTT 300

Db 399 ttaactttcaacttcgacttcgcagacagactatttctcgtaaggctcgaagaagcttggtc 458

QY 301 AGGATTTCACTTGAACAAAAGCAACCAAGCAAGAAATCTTGAACCTATTAATAA 360

Db 459 agcgatcagctagtagacaaaagcaaccaagcaagaatcttgacctataataataa 518

QY 361 GGTCTACATGCTCTAATGGAGACTATGGAAATGCACAGACACTCAAAATCTATGGTAA 420

Db 519 ggtctacatgctctaaatgggaactatggaaatgcagacagactcaaaaactactatgttaa 578

QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCCTTGCTGGCTGGAAATGCTTCAGGCACC 480

Db 579 agacctcaataatttaagtttaacctagtttagccttgcttgcttggaatgctcgaagcacc 638

QY 481 AAACCAATATGACCCCTTATTTACATCCAGAAAGCAGCCCAACACCGCGGAATCTTGCTT 540

Db 639 aaaccaatattgaccaccttaacataccagaagcagcccaagaccgcggaacttggtctt 698

QY 541 ATCTGAATGAATAATGAGGCTCATCTCTGTCGTAACAGATGAGAAAGCAGTCAATAC 600

Db 699 atctgaatgaaataatgagcctacacccctgcgcggaagaatgagaagaagccaataac 758

QY 601 ACCAATTAATGAGTACATCAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 660

Db 759 accaatattcgtatggtctacaagaatctcaaaatcagaagaatttaaccttgcttaccatgga 818

QY 661 TAATTAATCTCAAGAGTATCAATCAATCAAGTTGAAGAAAGAAACAGGCTTAATCACTAC 720

Db 819 taattacctcaaggaagctatcaatacaagttagaagaagaacaggtctataacctactcac 878

QY 721 AACGGGATGATGCTCAATCAAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 780

Db 879 aacgggattgattgcttcaacaataatgtagaccaagaagctcaaaaactcgttggtatc 938

QY 781 TTACAATACAGACCAATACGTTGCTTATCCAGAGATGATTAATGCAAGTCTGCTTACAT 840

Db 939 ttacaatacagagaatatacgttgcctaccagaagatgaaatgtaagtcgtcttaccat 998

QY 841 TGTGATGTTTCTAAGGTTAAGTCAATTTGCCACTAGAGACAGCATAGTAAAGTAA 900

Db 999 tgtgatgatttctaaagctaaagctatgcctcagctaggaagcagcatcagtaagtaa 1058

QY 901 TGTTCCTTGGGATTTAACCAAGCAGTAGAACAACCGGACGTGGGATCAATATGAA 960

Db 1059 tgttccttgggaatttaaccagaagcagtagaacaacacgcgagctgggatacctaataatgaa 1118

QY 961 ACCGATACAGACATCTATCTCTGCTTGGAGTACGGTGTCTACAGTTCAACGTCTACTAT 1020

Db 1119 accgatcacagacatctatctcttgcttggttagctgagtcagtcagtcagtcagtcactat 1178

QY 1021 CGTTCAGATGATGACCTTATACCTACCTGGGACAAATATCTCTTTATTAATGGGATAG 1080

Db 1179 cgttcagatgtagccctatactacccttggtgacaaataccctgtttataacttggagatg 1238

QY 1081 GGGCTACTTGGCAACATCACTTGCATATACGCCCTGCAACATCGGGAACGTCGCCAGC 1140

Db 1239 gggctacttggcaacatcaacttgcataacagccctgcacaacatcgcgaaaagctcccaagc 1298

QY 1141 CGTGAATACTTAAACAAGTTCGAGCTCAACCGGCCCAAGACTTTCTTAATGGTCTAGC 1200

Db 1299 cgtgaaactcttaacaagaagtcgagcaaccaacgcgcaagcttctctaattggtctcgg 1358

QY 1201 AATGACTACCCAGTATTTCTACTACTCAATATGCAATTTCAAGTAAACCAACGAATAGA 1260

Db 1359 aatgactaccacaagatttcaactactcaaatgccaatttcaagtaaacacaacagataga 1418

QY 1261 CAATAAATATGAGACAGTGTGAAAGATGAGGCTGAGCTTACGCTGCTTGGCAATGG 1320

Db 1419 caaataaataatgagcaagtagtgaagaatgagtcgagcttgaagcttgccttgcaaatgg 1478

QY 1321 TGAACCTTACTATTAACCAATGTATATCAATAAAGTGTCTTTAGTATGAGAGTGAATA 1380

Db 1479 tgaacttactataaacaacatgataatcaataaagtcgtctttagtgatggagtgaaaa 1538

QY 1381 AGAGTTCTTAAATGTCGAACTGTCGCAATGAAGAAAGCAACGCTTAATATGATGACCGA 1440

Db 1539 agagttctctaattcggaactcgtgcataatgaagaagaagcagcctatatatgatacgcga 1598

```

Oy 1441 CATGATGAAAACAGCTCTTACTTATGGAACAGGAGCAATGCTATCTTGGTTGGCTCCC 1500
Db 1599 CATGATGAAAACAGCTCTTACTTATGGAACAGGAGCAATGCTATCTTGGTTGGCTCCC 1658
Oy 1501 TCAGGCTGTGTAACAGCAACCTTAACTATATAGACAGCAAGAAATTGAAACACATCAA 1560
Db 1659 TCAGGCTGTGTAACAGCAACCTTAACTATATAGACAGCAAGAAATTGAAACACATCAA 1718
Oy 1561 GACCTTCGAATTTGTAGCACTGATGAACATTTTGTGGCTATACGCGTAAATATTCAAT 1620
Db 1719 GACCTTCGAATTTGTAGCACTGATGAACATTTTGTGGCTATACGCGTAAATATTCAAT 1778
Oy 1621 GGCCTATATGAGCAGGCTTCTTCTAACCGTGTAGACACACACTTGTAGGCAATGGCGTTAGCGT 1680
Db 1779 GGCCTATATGAGCAGGCTTCTTCTAACCGTGTAGACACACACTTGTAGGCAATGGCGTTAGCGT 1838
Oy 1681 CGTGCCCAAGTTTACCGCTCTATGATGATGACCTTACCTGTCTGTAAAGGAACATTCAGAGA 1740
Db 1839 CGTGCCCAAGTTTACCGCTCTATGATGATGACCTTACCTGTCTGTAAAGGAACATTCAGAGA 1898
Oy 1741 TTGGCAATATACCAAGGGGCTCTTACAGAAATGGAGAAATTCGATTTAAAAATGGTCTCG 1800
Db 1899 TTGGCAATATACCAAGGGGCTCTTACAGAAATGGAGAAATTCGATTTAAAAATGGTCTCG 1958
Oy 1801 TTCTACGAGGAACCTACCTGCTCCGACAAACACCCCATCATCACTGGAAGTTGAAGTCAATC 1860
Db 1959 TTCTACGAGGAACCTACCTGCTCCGACAAACACCCCATCATCACTGGAAGTTGAAGTCAATC 2018
Oy 1861 ATCAGATAGTTCACACTTCACAGTCTAGCTCAACCACTCAAGCAACAATTTATAGTACGAC 1920
Db 2019 ATCAGATAGTTCACACTTCACAGTCTAGCTCAACCACTCAAGCAACAATTTATAGTACGAC 2078
Oy 1921 TACCAATCTTAACATATATACGCAACATCAAAATGCAACCCCTGATCAACAAATACAGAA 1980
Db 2079 TACCAATCTTAACATATATACGCAACATCAAAATGCAACCCCTGATCAACAAATACAGAA 2138
Oy 1981 TCCTCAACCAACACACACCA 1999
Db 2139 TCCTCAACCAACACACACCA 2157

RESULT 10
US-10-072-851-9325
: Sequence 9325, Application US/10072851
: GENERAL INFORMATION:
: APPLICANT: Carr, Grant J.
: APPLICANT: Xu, H. Howard
: APPLICANT: Foulkes, J. Gordon
: APPLICANT: Zamudio, Carlos
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Tirawick, John D.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Roemer, Terry
: APPLICANT: Jiang, Bo
: APPLICANT: Boone, Charles
: APPLICANT: Bussey, Howard
: TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
: FILE REFERENCE: ELITRA.028A
: CURRENT APPLICATION NUMBER: US/10/072,851
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: 60/267,636
: NUMBER OF SEQ ID NOS: 15811
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9325
: LENGTH: 2160
: TYPE: DNA
: ORGANISM: Streptococcus pneumoniae

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2160)
US-10-072-851-9325

```

Query Match	99.1%;	Score 1981.4;	DB 37;	Length 2160;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1988; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

OY	1	TAATAATGTACGACAATTA	AAATTAATCACTGCTATGTCGCTTGGCTTGTAACGGCCGGCTAA	60
Db	159	taaaatctacagacaataa	aaataactcaatctgcagcttgggtcttgcgaacggccggctaa	218
OY	61	TGCCCAAGCTAAATGATAT	TTCGCCACAGATTTGGTTAAGGCAATCGTTTCTATGAGAGACCA	120
Db	219	tggcccaagctaaatgatatt	tcccacagatttggtttaagaaatcgtttctatcogaagacca	278
OY	121	TGCGTTCTTGACCA	CCAGGGGGAATTGATACCATCGTATCTGGGAGCTTCTTGCGCAA	180
Db	279	tgcgtctcttcagacaca	aggggagatgatacaccatcgtatccctggagcttcttcgcgaa	338
OY	181	TCGCGCAAGAATTTCC	TCCCAAGTGGATCACTGTCACCCCAACAAGTTGATTAATCTTAC	240
Db	339	tcgtcgaaagaatctcc	tcccaaggatgatacgccttcacatacaagcttgataaagttgac	398
OY	241	TTACTTTTCAACTTTC	GCATCCGACGAGACTATTTTCTCTAAGGCGTCAAGAACTTGGSTT	300
Db	399	ttaactttcaactcttgac	tttcgcagactatcttcctgtaaggtctcaggaaagcttggtt	458
OY	301	AGCGATTCAAGTTAGAA	CAAAAAAGCAACGACGAAGAAATCTTGACCTACTATATAATAA	360
Db	459	agcgattcagcttaga	acaaagaacaacgaagaatcttgcactactatataataaa	518
OY	361	GGTGTCAATGTCTAA	TGAGGGAACATGTGAGATTCGACAGAGGTCGTAACACTCTATAGTAA	420
Db	519	ggcttaacaatgctaa	tgggaactatggaatgcgaagaagactcaaaactactaatgtaa	578
OY	421	AGACCTCAATTAATTT	TAAGTTTACCTCAGTTAGCGCTTGGCTGGAGAAATGCCAGGCAAC	480
Db	579	agacctcaataat	ttaagttaacttaagttacctaagttagcttgcgtgttgaaatgcttcagacc	638
OY	481	AAACCAATATGACCC	CCCTATTCACATCCAGAGAGAGCCCAAGACGCGCAAACTTGCTT	540
Db	639	aaaccaata	tgacccattatcacatccagaagaagcccaagccgcggaacttgctt	698
OY	541	ATCTGAATGAAAAAT	CAAGGCTACTCTCTGTGAACAGTATGAGAAAGCATCAATAC	600
Db	699	atctgaaatgaaaaa	atcaagaagctacatctctctgtaacagatgagaagaagctaaatac	758
OY	601	ACCAATTAAGTGA	GGACTACAAAGTCTCAAAATCAGACAAGTAATTAACCTGGCTTACATGGA	660
Db	759	accaatctacgata	ggggtctacaagaagttccaatcaagaagaagtaattacacctgtcttacaatga	818
OY	661	TAAATFACCTCAAG	AAGATCTCATCAATCAAGTTGAGAGAAGCAAGGCTATTAACCTACTCAC	720
Db	819	taattacctcaag	aagatctcaatcaagttgaagaagaacagctataacctactcac	878
OY	721	AACCTGGATGGA	TGTCTACACAAATGTAAACCAAGAAGTCAAAAAACTCTGTGGGATAT	780
Db	879	aacctggagatgata	gtcttacaacaaatgtaagacaagaagttcaaaaacatctgttggatatac	938
OY	781	TTACAAATACAGAG	AATAGCTTGCTCATTCAGACGATGTAATTCGAAGTCGGCTTACCAT	840
Db	939	ttacaatacaga	agaataagttgtccataccagaagatgtgcaagtcgcttctaccat	998
OY	841	TGTTGATGTTTCA	AGCGTAAAGTCAATGCCAGCTAGGAGCAGCCATCAATCAAGTAA	900
Db	999	tgttgaatgtttca	acgtaagaagttatgcccagctagaagacgcatcagtaagtaa	1058
OY	901	TGTTTCTTGGAA	TTAATTAACCAAGCACTAGTAAACCAACGCGACTGGGATCAACTATGAA	960
Db	1059	tgtttccttcgtgaat	ttaacaaagctatgaacaaacccgagcttgggtatcaactatgaa	1118

OY	961	ACGATCAAGACATATCTCTCGCTTGGAGTACGGTGCTACAGTTGCACTGCTACTAT	1020
Db	1119	accgatacaagaactatgctccctgccttgagatcagtgctcaagtcgaactgccaatt	1178
OY	1021	CGTTACAGATGAGCCCTATATCTACCTGGGACAAATATCTCTGTTTATTACTGGGATG	1080
Db	1179	cgttacagatgagccctataactaccctcggacaatatcccttglttatactggtatg	1238
OY	1081	GGGCTACTTTTGACACATCTACCTTGCAATACGCTTCGACAACTGCGAAGCTCCAGC	1140
Db	1239	gggctactcttggcaacataccactctgaataacgcccctgcacaacatcggaacgctccagc	1298
OY	1141	CGTGGAAATCTTAACAAAGGTGCGAGCTCAACCGCGCAAGACTTTTCTTAATAGGTCTAG	1200
Db	1299	cgtygaacacttaacaagaagtcggaactcaacgcgccaaagacttccctaattgctcgg	1358
OY	1201	AATCGACTACCCAGATTTCTACTCTCAATATGCGATTTTCAAGTAAACACACCGAATCAGA	1260
Db	1359	aatcgactaccacaagtaattcaactaccaaatgcatcttaagtaacacaaccgaaatcaga	1418
OY	1281	CAAAAAATATGGACCAAGTGTGTAAGAAAGATGGGTGCTTGCGCTTGCGCTTTGGCAAATGG	1320
Db	1419	caaaaaatahygaacaagatgynaagaatggctgctgtcttaagcttgcccttgycnaayg	1478
OY	1321	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTAAAA	1380
Db	1479	tggaaacttaactataaacaatgataatcacaataagtcgctcttagtgatgagygaa	1538
OY	1381	AGATTCTCTAATGTGCGAATCGTGCCATGAAAGCAACAGACGCTATATGATGACCGA	1440
Db	1539	agattctcttaattgcygaactcgtgcataagaaagaaacgaagctataatgataccga	1598
OY	1441	CATGATNAACAAAGTCTTGACTATATGAACTGGACCAATGGCTATTTGGCTGGCTCC	1500
Db	1599	catgataaaacagtccttgagttatggaactggaogaatgctatcttgcttgctcc	1658
OY	1501	TCAGGCTGTAAACACAGAACTCTTACTATACAGACGAGAAATTGAAAAACACATCAA	1560
Db	1659	tcaggctgttaaaacagaaaccttaactatacagaacgagaatctgaaaaccacaatcaa	1718
OY	1561	GACCTCCATATTTGTACCACTGATGAACGATTTTGTGCTATACGGTAATATTCAT	1620
Db	1719	gaacctccaattctgtagaccggaatgaacattctgcgtctatacgcgtlaatatccaat	1778
OY	1621	GGCTGTATGACAGAGCTATTTCTTAACCGTGTGCACCACTTGTAGGCAATGGCTTACGGT	1680
Db	1779	ggctgtatgtagaagcgctattcttaaccgctctgaaaccaactctgtagcaatgaccttaagct	1838
OY	1661	CGCTGCCAAAGTTTACCGGCTCTATGATGACCTTACCTGCTGAAGAGACATCCAGAAGA	1740
Db	1839	cgcgcgcaaaagtttaaccgctctatgtagcttaaccctgctgaaagaaacatccaagaaga	1898
OY	1741	TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAAAATGGGCTCG	1800
Db	1899	ttggaatatataccagagggtctctaacgaaatgagaatctcgatttaaaaatgtygtccg	1958
OY	1801	TTTACGTGGAATCTACCTGCTCCACACAGACCCCATCACTGAAAGTCTCAAGCTCATC	1860
Db	1959	tttcaagcttgagcttcaacctctctcaacaacaaccccatcaactgaagaattcaagctc	2018
OY	1861	ATCGATATAGTTCAACCTTCAAGCTGTAGGTCACCACTCCACACCAAAATATATGATGAC	1920
Db	2019	atcgaatagttcaacttcaacagctctgactcaaacactccaagcacaataatagtaagac	2078
OY	1921	TACCAATCTCTAACATTAATACGCAACATCAATAACACCCCTGATCAACAAATCAGAA	1980
Db	2079	taccaatctcaacaataataatgcaacaacatcaataaacaccttgctcaacaataatcagaa	2138
OY	1981	TCCTCAACGACGACACCA 1999	
Db	2139	tcctcaaacgacgacaccaa 2157	

```

RESULT 11
US-60-061-998-596/c
; Sequence 596, Application US/60061998
; GENERAL INFORMATION:
; APPLICANT: LAGACE, ROBERT E.
; APPLICANT: CORLEY, NEIL C.
; APPLICANT: RUSSO, FRANK D.
; APPLICANT: HANN, AMY L.
; APPLICANT: HEATH, JOE D.
; APPLICANT: FINNEY, GREGORY L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/061,998
; FILING DATE:
;
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0006-2P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 596:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: SPN1C601
US-60-061-998-596

Query Match          99.1%; Score 1981.4; DB 45; Length 1033;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 198; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 TAAATCTACGACGATTAATAATCACTCATCTGCTGACTGGTTCTGAAGCGCGGTCAA 60
9718 TAAATCTACGACGATTAATAATCACTCATCTGCTGACTGGTTCTGAAGCGCGGTCAA 9659
61 TGCCCAAGCTAATGATATTCCACAGATTGGTTAAGGCATCGTTCTATCGAAGACCA 120
9658 TGCCCAAGCTAATGATATTCCACAGATTGGTTAAGGCATCGTTCTATCGAAGACCA 9599
121 TCGCTTCTTCGACCAAGGGGATGATACATCGGTATCCGAGAGCTTTCTCGGCA 180
9598 TCGCTTCTTCGACCAAGGGGATGATACATCGGTATCCGAGAGCTTTCTCGGCA 9539
181 TCTGCAAGCAATTCCTCCACAGGTGATCACTCCACCCAAACAGTTGATTAACTTGAC 240
9538 TCTGCAAGCAATTCCTCCACAGGTGATCACTCCACCCAAACAGTTGATTAACTTGAC 9479
241 TTACTTTTCACTTGACATCTCCGACCAAGCTATTCTCGTAAGGCTCAGGAAGCTTGGTT 300
9478 TTACTTTTCACTTGACATCTCCGACCAAGCTATTCTCGTAAGGCTCAGGAAGCTTGGTT 9419

```


ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PM-0009-2 P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-416
INFORMATION FOR SEO ID NO: 596:
SEQUENCE CHARACTERISTICS:
LENGTH: 9845 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: genomic DNA
IMMEDIATE SOURCE:
CLONE: SPN2C598
US-60-068-175-596

Query Match 81.6%; Score 1631.2; DB 45; Length 9845;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1776; Conservative 0; Mismatches 223; Indels 1; Gaps 1;

QY 1 TAAATCTACGACATAAAAAATCACTCATCTGACTTGGGTTCTGAACGCCGCTCAA 60
DB 9734 TAAATCTACGACATAAAAAATCACTCATCTGACTTGGGTTCTGAACGCCGCTCAA 9675
QY 61 TGCCCAAGCTAATATATATTTCCACAGATTGGTTAAGCAATCGTTTCTATCAGAACCA 120
DB 9674 TGCCCAAGCTAATATATTTCCACAGATTGGTTAAGCAATCGTTTCTATCAGAACCA 9615
QY 121 TCGCTTCTGCACACAGAGGGGATTTACCATCCGTATCCCTGGAGGCTTTTCGCCAA 180
DB 9614 TCGCTTCTGCACACAGAGGGGATTTACCATCCGTATCCCTGGAGGCTTTTCGCCAA 9555
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCACCCACAGTTGATTAAAGTTAC 240
DB 9554 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCACCCACAGTTGATTAAAGTTAC 9495
QY 241 TTACTTTTCAACTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 300
DB 9494 TTACTTTTCAACTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 9435
QY 301 AGCGATTAGTTAGACAAAAAGCAACCAAGCAAGAAATCTTGAACCTATATATAATAA 360
DB 9434 AGCGATTAGTTAGACAAAAAGCAACCAAGCAAGAAATCTTGAACCTATATATAATAA 9375
QY 361 GGTCTACATGTCTAATGGAAGTATGGAATGACAGACAGACTCAAACTACTATGTGTA 420
DB 9374 GGTCTACATGTCTAATGGAAGTATGGAATGACAGACAGACTCAAACTACTATGTGTA 9315
QY 421 AGACCTCAATATTTAAGTTTACTCTAGTTAGCTTGGCTGGTGAAGTGGCTCAGGGACC 480
DB 9314 AGACCTCAATATTTAAGTTTACTCTAGTTAGCTTGGCTGGTGAAGTGGCTCAGGGACC 9255
QY 481 AAACCAATATGACCCCTATTTCATCCAGAGAGCCCAAGCCGCGAAATCTTGCTTT 540
DB 9254 AAACCAATATGACCCCTATTTCATCCAGAGAGCCCAAGCCGCGAAATCTTGCTTT 9195
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAAGCATATAC 600
DB 9194 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAAGCATATAC 9135
QY 601 ACCAATTAAGTGGAGCTACAAAGTCCAAATGCAAGTAAATTAACCT-GCTACATAGG 659
DB 9134 ACCAATTAAGTGGAGCTACAAAGTCCAAATGCAAGTAAATTAACCTGCTACATAGG 9075
QY 660 ATAATTACTCAAGAGTATCATCAATCAAGTTGAGAGAAAGCAAGCTATTAACCTACTCA 719
DB 9074 ATAATTACTCAAGAGTATCATCAATCAAGTTGAGAGAAAGCAAGCTATTAACCTACTCA 9015
QY 720 CAATGGAGATGATGTCTACACAAATGTAGACCAAGAAAGCTCAAAAACATCTGGGGATA 779
DB 720 CAATGGAGATGATGTCTACACAAATGTAGACCAAGAAAGCTCAAAAACATCTGGGGATA 779

DB 9014 CTACTGGAGATGATGTTTACACAAATGTAGACCAAGAAAGCTCAAAAACATCTGGGGATA 8955
QY 780 TTACAAATACAGCAATATAGCTGCTATCCAGAGATGAATGCAAGTCCGCTTACCA 839
DB 8954 TCTACAACTCCGATCAATAGCTCTTACCGTACAGATGATTTGCAAGTCCGATCAGG 8895
QY 840 TTGTGATGTTTCAAGGTAAGTCAATTCGCCAGTACAGACGCCATCAGTCAAGTA 899
DB 8894 TCGTAGATGTTTCAAAATGCTAAAGTCAATCCGCCAACTGGAGCTGCACCAACAGTA 8835
QY 900 ATGTTTCTCGGAATTAACCAACAGTACGAACAAACCCGAGTGGGATCACTATGA 959
DB 8834 ACGTTTATTTGTACCAACCAACCTGTGGAAACCAATCGAGCTGGGTTCTCTATGA 8775
QY 960 AACCATACAGACTATGCTCTGCTGAGTACGAGTGTCTAGATTCACCTACTACTA 1019
DB 8774 AACCATACAGACTATGCTCTGCTGAGTACGAGTGTCTAGATTCACCTACTACTA 8715
QY 1020 TCGTTACAGATGAGCCCTATATACCTCGGACAAATCTCTGTTTAACTGGGATA 1079
DB 8714 TCGTTAATGATATCTTATATATCCGGAGACAGACACTGTCTACAACTGGGATA 8655
QY 1080 GGGCTACTTTGGCAATCACCCTTGCAATACGCCCTGCAACATTCGGAACGCTCCAG 1139
DB 8654 GAGCATATTTCTGTAATATTAATCTGCAATATGCTTCAACAAATCAAGAAATGTCACAG 8595
QY 1140 CCGTGSAACTCTAACAAGGTGGAGCTCAACGCCGCAACACTTCTCTAAATGGCTAG 1199
DB 8594 CCGTTGAGACTTTGAATAGGCTGCTAGATGATGATGATTAACCTCTCTATATGCTTG 8535
QY 1200 GAATGACTACCAAGTATCTACTACTCAATTCGATTTCAAGTACACCAACCAATCAG 1259
DB 8534 GTATGACTATTCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8475
QY 1260 ACAAAAAATATGAGCAAGTATGTAAGTATGTAAGTATGTAAGTATGTAAGTATGTAAG 1319
DB 8474 ATAAACATATAGGCAAGTATGTAAGTATGTAAGTATGTAAGTATGTAAGTATGTAAG 8415
QY 1320 GTGACCTACTATTAACCAATGATATATCAATTAAGTATGTAAGTATGTAAGTATGTAAG 1379
DB 8414 GTGACCTACTATTAACCAATGATATATCAATTAAGTATGTAAGTATGTAAGTATGTAAG 8355
QY 1380 AAGATTTCTTAATGTGGAAGTCTGTCATGATGAAGCAAGCAAGCTATATATGATGACG 1439
DB 8354 AAGATTTCTTAATGTGGAAGTCTGTCATGATGAAGCAAGCAAGCTATATATGATGACG 8295
QY 1440 ACATGATGAAGCAAGTCTGTCATGATGAAGCAAGCAAGCTATATATGATGATGACG 1499
DB 8294 ACATGATGAAGCAAGTCTGTCATGATGAAGCAAGCAAGCTATATATGATGATGACG 8235
QY 1500 CTCAGGCTGGTAAAGCAAGCAAGCTCTAATCTATACAGAGCAAGCAAGCTATATATGATG 1559
DB 8234 CTCAGGCTGGTAAAGCAAGCAAGCTCTAATCTATACAGAGCAAGCAAGCTATATATGATG 8175
QY 1560 AGACCTCTCAATTTGTAGCAAGCTGATGAAGTATGCTGCTATAGCGTAAATATTTCAA 1619
DB 8174 AGACCTCTCAATTTGTAGCAAGCTGATGAAGTATGCTGCTATAGCGTAAATATTTCAA 8115
QY 1620 TGGCTGATGAGCAAGGCTATTTCTAAGCTGTGACACACACTGTGAGCAATGGCTTACG 1679
DB 8114 TGGCTGATGAGCAAGGCTATTTCTAAGCTGTGACACACACTGTGAGCAATGGCTTACG 8055
QY 1680 TCGCTGCAAGTATACCGCTATGATGAGCTCTGCTGGAAGCAAGCAAGTACCAAG 1739
DB 8054 TCGCTGCAAGTATACCGCTATGATGAGCTCTGCTGGAAGCAAGCAAGTACCAAG 1739
QY 1740 ATTGGAATATACAGAGAGGCTCTACAGAAATGAGAAATTCGATTTAAATATGCTCTC 1799
DB 7994 ACTGAGAGATGCCAGAGGACTTTTCAAGAAAGGGAATTTGATTCAAAATATGAGACTC 7935
QY 1800 GTTCTACGTTGAACTCACTGCTCTCAGACACACCCCATCACTAGCAAGTCAAGCTCAT 1859
DB 7934 GCCCAATATGAGTGAACCCCTTACTCAACAAATCTCAACAGCTGAAGTCAAGCTCAT 7875

OY 1860 CATCAGATAGTTCACAGTCTACAGTCAACCACTCCAGCACAATATAGTACGA 1919
|||||
Db 7874 CATCAGATAGTTCACAGTCTACAGTCTACAGCACAATATATAGTACGA 7815
OY 1920 CTACCAATCTTAACAATTAATAGCAACAATCAAAATCAACCCCTGATCAGAAAATCAGA 1979
|||||
Db 7814 CTACCAATCTTAACAATTAATAGCAACAATCAAAATCAACCCCTGATCAGAAAATCAGA 7755
OY 1980 ATCTCAACCAAGCAGACCA 1999
|||||
Db 7754 ATCTCAACCAAGCAGACCA 7735

RESULT 13

US-08-116-541-3
: Sequence 3, Application US/08116541
: GENERAL INFORMATION:
: APPLICANT: Masure, H. Robert
: APPLICANT: Pearce, Barbara J.
: APPLICANT: Toumanen, Elaine
: TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
: TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/116, 541
: FILING DATE: 19930901
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-069
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
: TELEX: 133521
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 960 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Streptococcus pneumoniae
: STRAIN: R6
: IMMEDIATE SOURCE:
: CLONE: SPRU42
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..960
: US-08-116-541-3

Query Match 47.2% Score 944.2; DB 5; Length 960;
Best Local Similarity 99.7% Pred. No. 1e-256;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TAAATCTACGACATAAATAATCAACTGCTGCTGCTGGTCTGGAACCGCGCTCAA 60

Db 12 TAAATCTACGACATAAATAATCAACTGCTGCTGCTGGTCTGGAACCGCGCTCAA 71
OY 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGCAATCGTTCTATGGAAGACA 120
Db 72 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGCAATCGTTCTATGGAAGACA 131
OY 121 TCGCTTCTTCGACACAGAGGGGATGATACATCGGTATCCGGAGGCTTCTTCGCA 180
Db 132 TCGCTTCTTCGACACAGAGGGGATGATACATCGGTATCCGGAGGCTTCTTCGCA 191
OY 181 TCTGCAAGCAATTCCTCCAAAGTGATCACTCCACCCAGTGTATTAAGTTGAC 240
Db 192 TCTGCAAGCAATTCCTCCAAAGTGATCACTCCACCCAGTGTATTAAGTTGAC 251
OY 241 TTACTTTTCAACTTCGACTCCGACACACTATTTCTGTAAAGCTCAGAAAGCTTGGT 300
Db 252 TTACTTTTCAACTTCGACTCCGACACACTATTTCTGTAAAGCTCAGAAAGCTTGGT 311
OY 301 AGCGATCAGTTAGAAACAAACCAAGCAAGAAATCTTGACCTATATATATA 360
Db 312 AGCGATCAGTTAGAAACAAACCAAGCAAGAAATCTTGACCTATATATATA 371
OY 361 GGTCTACATGCTTAATGGGAATGATGATGACAGACAGCTCAAACTACTATGTAA 420
Db 372 GGTCTACATGCTTAATGGGAATGATGATGACAGACAGCTCAAACTACTATGTAA 431
OY 421 AGACCTCAATATATTAAGTTACCTGACTAGCTTCTGCTGGTGAATGCTCAGGACC 480
Db 432 AGACCTCAATATATTAAGTTACCTGACTAGCTTCTGCTGGTGAATGCTCAGGACC 491
OY 481 AAACCAATATGACCCCTATTCATCCAGAGCAGCCCAAGCCGCAAAATTTGGCTT 540
Db 492 AAACCAATATGACCCCTATTCATCCAGAGCAGCCCAAGCCGCAAAATTTGGCTT 551
OY 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGACAGTATGAGAAAGCAGTCAATAC 600
Db 552 ATCTGAATGAAAAATCAAGGCTACATCTGCTGACAGTATGAGAAAGCAGTCAATAC 611
OY 601 ACCAATTAAGTGAAGTACCAAAAGTCTCAAAATCAGCAAGTATTAATCCCTGCTTACATGA 660
Db 612 ACCAATTAAGTGAAGTACCAAAAGTCTCAAAATCAGCAAGTATTAATCCCTGCTTACATGA 671
OY 661 TAATTACTCAAGGAAGTATCAATCAAGTGAAGAAAGAGGCTTAATCACTACTAC 720
Db 672 TAATTACTCAAGGAAGTATCAATCAAGTGAAGAAAGAGGCTTAATCACTACTAC 731
OY 721 AACTGGATGATGCTCTACAAATGTAGACCAAGCTCAAAACATCTGTGGATAT 780
Db 732 AACTGGATGATGCTCTACAAATGTAGACCAAGCTCAAAACATCTGTGGATAT 791
OY 781 TTACATACAGAGCAATAGTTCCTATCCAGACAGATGAATTCAGTGGCTTACAT 840
Db 792 TTACATACAGAGCAATAGTTCCTATCCAGACAGATGAATTCAGTGGCTTACAT 851
OY 841 TGTGATGTTTCTAAGCTAAAGTCAATGCCAGCTAGAGCAGCCATCGATCAAGTAA 900
Db 852 TGTGATGTTTCTAAGCTAAAGTCAATGCCAGCTAGAGCAGCCATCGATCAAGTAA 911
OY 901 TGTTCCTTCGGAATTAACCAAGAGTAAACCAACCGCGCTGGGA 949
Db 912 TGTTCCTTCGGAATTAACCAAGAGTAAACCAACCGCGCTGGGA 960

RESULT 14

PCT-US97-14436-139
: Sequence 139, Application PC/TUS9714436
: GENERAL INFORMATION:
: APPLICANT: Black, Michael
: APPLICANT: Hodgson, John
: APPLICANT: Knowles, David
: APPLICANT: Lonetto, Michael
: APPLICANT: Nicholas, Richard

APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
POLYPEPTIDES AND THEIR USES
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
PCT-US97-14436-139

Query Match 40.58; Score 808.6; DB 1; Length 2172;
Best Local Similarly 99.4%; Pred. No. 5.1e-218;
Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1174 CGCGAAGCTTCTCTAATGCTAGGATCGACTACCCAGATTCACCTCAATGCG 1233
1 CGCGAAGCTTCTCTAATGCTAGGATCGACTACCCAGATTCACCTCAATGCG 60
DB 1234 CATTCAAGTAACACACGCAATCAGACAAAAATATGAGCAAGTAGTGAAGATGCG 1293
61 CATTCAAGTAACACACGCAATCAGACAAAAATATGAGCAAGTAGTGAAGATGCG 120
QY 1294 TGCCTGCTTACGCTCCCTTTGCAAAATGCTGCACTTACTATTAACCAATGATATCAATAA 1353
121 TGCCTGCTTACGCTCCCTTTGCAAAATGCTGCACTTACTATTAACCAATGATATCAATAA 180
DB 1354 AGTGTCTTAGTATGATGAGGAGTGAAGAAAGTCTCTAATGCTGGAAGTGTGCGCATGAA 1413
181 AGTGTCTTAGTATGATGAGGAGTGAAGAAAGTCTCTAATGCTGGAAGTGTGCGCATGAA 240
QY 1414 GAAACGACAGCCCTATATGATGACGACATGATGAAAAAGTCTTGAATATGAGACTGCG 1473
241 GAAACGACAGCCCTATATGATGACGACATGATGAAAAAGTCTTGAATATGAGACTGCG 300
DB 1474 ACG-AAATGCTATCTGCTTGGCTCCCTCAAGCTGTGTAAGCAAGAACTCTAATATA 1552
301 ACGAAATGCTATCTGCTTGGCTCCCTCAAGCTGTGTAAGCAAGAACTCTAATATA 360
QY 1533 CAGACGAGGAATTTGAACACATCAAGACCTCTCAATTTGATGACACCTGATGAATAT 1592
361 CAGACGAGGAATTTGAACACATCAAGACCTCTCAATTTGATGACACCTGATGAATAT 420
DB

QY 1593 TTGCTGCTATACGCGTAATATTCATGCTGTATGAGACAGGCTATTTCAACGCTCGA 1652
421 TTGCTGCTATACGCGTAATATTCATGCTGTATGAGACAGGCTATTTCAACGCTCGA 480
QY 1653 CACCCTTTGATGCAATGGCTTACGGTCTGCGCAAGTTTACCGCTATGATGACT 1712
481 CACCCTTTGATGCAATGGCTTACGGTCTGCGCAAGTTTACCGCTATGATGACT 540
DB 1713 ACCTGTGGAAGGAAGCAATCCAGAGATTTGGAATATACGAGGGGCTCTACAGAAATG 1772
541 ACCTGTGGAAGGAAGCAATCCAGAGATTTGGAATATACGAGGGGCTCTACAGAAATG 600
QY 1773 GAGAAATCGATTTTAAATGCTGCTGTCTACGTGGAACCTACCTCCACAAAC 1832
601 GAGAAATCGATTTTAAATGCTGCTGTCTACGTGGAACCTACCTCCACAAAC 660
DB 1833 CCCCATCACTGAAAGTTTCAAGCTCATCATGATGATTTCACTTACAGCTTACGCTCAA 1892
661 CCCCATCACTGAAAGTTTCAAGCTCATCATGATGATTTCACTTACAGCTTACGCTCAA 720
QY 1893 CCACCTCAAGCAAAATTAATAGTACGACTACCAATCCTAACAAATATACGCAATCAA 1952
721 CCACCTCAAGCAAAATTAATAGTACGACTACCAATCCTAACAAATATACGCAATCAA 780
DB 1953 ATACAAACCCCTGATCAACAAATCAGAAATCCTCAACGACGACAAACA 1999
781 ATACAAACCCCTGATCAACAAATCAGAAATCCTCAACGACGACAAACA 827
DB

RESULT 15
US-08-911-503-139
Sequence 139, Application US/08911503
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
POLYPEPTIDES AND THEIR USES
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,503
FILING DATE: 15-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-911-503-139

Query Match 40.5%; Score 808.6; DB 13; Length 2172;
Best Local Similarity 99.4%; Pred. No. 5,1e-218;
Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

OY 1174 CGCCAGACTTCTCTAAATGCTAGGAATCGACTACCCAGTATCTACTCTCAATGC 1233
|||||
DB 1 CGCCAGACTTCTCTAAATGCTAGGAATCGACTACCCAGTATCTACTCTCAATGC 60
OY 1234 CATTTCAGTAAACACACACCGAATTCAGACAAAAATATGAGCAGTGTGAAGATGCG 1293
|||||
DB 61 CATTTCAGTAAACACACCGAATTCAGACAAAAATATGAGCAGTGTGAAGATGCG 120
OY 1294 TGCCTGCTTACGCTGCTTGGCAATGCTGAGACTTACTTAACCATGTATATCCATTA 1353
|||||
DB 121 TGCCTGCTTACGCTGCTTGGCAATGCTGAGACTTACTTAACCATGTATATCCATTA 180
OY 1354 AGTCGCTTTTATGATGAGGAGTGAAGAAAGTTCTTAATGTCGGAACCTGTCATGAA 1413
|||||
DB 181 AGTCGCTTTTATGATGAGGAGTGAAGAAAGTTCTTAATGTCGGAACCTGTCATGAA 240
OY 1414 GGAACGACAGCCTATATGATGACCGACATGATGAAAAACAGTCTTGACTTATGAACTGG 1473
|||||
DB 241 GGAACGACAGCCTATATGATGACCGACATGATGAAAAACAGTCTTGACTTATGAACTGG 300
OY 1474 ACG-AAATGCTATCTGTGCTGGCTCCCTCAGGCTGTGTAACAGGAACCTTAACATA 1532
|||||
DB 301 ACGAAATGCTATCTGTGCTGGCTCCCTCAGGCTGTGTAACAGGAACCTTAACATA 360
OY 1533 CAGACGAGGAATTTGAAAAACACATCAAGACCTCTCAATTTGAGACCTGATGAACAT 1592
|||||
DB 361 CAGACGAGGAATTTGAAAAACACATCAAGACCTCTCTCAATTTGAGACCTGATGAACAT 420
OY 1593 TTGCTGGCTATACGCGTAAATATTCATGCTGTATGACAGGCTATTTAAACGCTTGA 1652
|||||
DB 421 TTGCTGGCTATACGCGTAAATATTCATGCTGTATGACAGGCTATTTAAACGCTTGA 480
OY 1653 CACCACTTGTAGGCAATGGCTTACGGTGGCTGCCAAATTTACCGCTCTATGATGACCT 1712
|||||
DB 481 CACCACTTGTAGGCAATGGCTTACGGTGGCTGCCAAATTTACCGCTCTATGATGACCT 540
OY 1713 ACCTGCTGAAGAGCAATTCAGAAATTTGAATATACAGAGGGGCTCTACAGAAATG 1772
|||||
DB 541 ACCTGCTGAAGAGCAATTCAGAAATTTGAATATATACAGAGGGGCTCTACAGAAATG 600
OY 1773 GAGAAATTCGTAATTTAAAAATGGTGTCTGTCTACGTGAATCACCCTGCTCCACAACAAC 1832
|||||
DB 601 GAGAAATTCGTAATTTAAAAATGGTGTCTGTCTACGTGAATCACCCTGCTCCACAACAAC 660
OY 1833 CCCCATCAACTGAAAGTTCACTCATCATGATAGTTCAACTTACAGTCTAGCTCAA 1892
|||||
DB 661 CCCCATCAACTGAAAGTTCACTCATCATGATAGTTCAACTTACAGTCTAGCTCAA 720
OY 1893 CCACATCAAGCAAAATATATAGAGACTTACCAATCTTAACAATATATAGCAACATCAA 1952
|||||
DB 721 CCACATCAAGCAAAATATATAGAGACTTACCAATCTTAACAATATATAGCAACATCAA 780
OY 1953 ATACAACCCCTGATCAACAAATCAGAAATCTCAACACAGACACACCA 1999
|||||
DB 781 ATACAACCCCTGATCAACAAATCAGAAATCTCAACACAGACACACCA 827

Search completed: August 11, 2002, 04:16:37
Job time: 9723 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2002, 02:09:49 ; Search time 442.87 seconds
(without alignments)
11011.969 Million cell updates/sec

Title: US-08-961-083-1
1999
Sequence: 1 TAAATCTACGACAAATAAAA.....ATCGTCACACGACACACCA 1999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1437365 seqs, 1219827628 residues
Total number of hits satisfying chosen parameters: 2874730

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	4	US-08-961-083-1
2	1989	100.0	10711	7	US-10-158-844-145
3	117.2	5.9	9100	7	US-10-121-120-27
4	59.2	3.0	2370	5	US-09-540-209B-3719
5	57.8	2.9	7571	7	US-10-158-844-2
6	55.2	2.8	428	5	US-09-673-476-206
7	51.8	2.6	363	5	US-09-673-476-626
8	44.4	2.2	1104	7	US-10-179-131-3144
9	44.2	2.2	1350	7	US-10-179-131-2857
10	40.6	2.0	223	5	US-09-673-476-438
11	40.4	2.0	1611	7	US-10-179-131-3015
12	40	2.0	610	7	US-10-027-632-5967
13	39.2	2.0	867	5	US-09-216-393B-340
14	39.2	2.0	867	5	US-09-216-393B-342
15	39.2	2.0	1397	5	US-09-216-393B-343
16	39.2	2.0	1397	5	US-09-216-393B-345
17	38.6	1.9	1143	7	US-10-179-131-1548
18	38	1.9	430	7	US-10-027-632-40693
19	37.6	1.9	2805	7	US-10-155-881-23712
20	37.4	1.9	242	5	US-09-789-189-1228
21	37.4	1.9	2034	7	US-10-179-131-3855
22	37.2	1.9	1611	8	US-60-360-039-46474
23	36.6	1.8	679	7	US-10-027-632-213202
24	36.6	1.8	2409	8	US-60-360-039-46149
25	36.6	1.8	8093	7	US-10-172-086-26

26	36.4	1.8	473	5	US-09-673-476-671	Sequence 671, App
27	36.4	1.8	9021	7	US-10-172-086-54	Sequence 54, Appl
28	36	1.8	689	7	US-10-027-632-148124	Sequence 148124, Appl
29	35.8	1.8	348	5	US-09-673-476-99	Sequence 99, Appl
30	35.6	1.8	445	5	US-09-918-995-15709	Sequence 15709, A
31	35.6	1.8	2373	5	US-09-540-209B-4136	Sequence 4136, Ap
32	35.4	1.8	8045	5	US-09-913-878A-1	Sequence 1, Appl1
33	35.2	1.8	1812	7	US-10-179-131-2414	Sequence 2414, Ap
34	35.2	1.8	3234	8	US-60-385-568-185	Sequence 185, App
35	35.2	1.8	3402	7	US-10-179-131-1603	Sequence 1603, Ap
36	35	1.8	510	7	US-10-027-632-82144	Sequence 82144, A
37	35	1.8	132205	5	US-09-578-519B-25	Sequence 309095,
38	35	1.8	132205	5	US-09-578-243B-25	Sequence 25, Appl
39	35	1.8	132205	5	US-09-578-518B-25	Sequence 25, Appl
40	35	1.8	132205	5	US-09-578-518B-25	Sequence 25, Appl
41	35	1.8	132205	5	US-09-578-518B-25	Sequence 25, Appl
42	35	1.8	132205	7	US-10-185-047-25	Sequence 25, Appl
43	35	1.8	132205	7	US-10-185-047A-25	Sequence 25, Appl
44	34.8	1.7	1655	7	US-10-175-525-86	Sequence 86, Appl
45	34.8	1.7	2709	7	US-10-179-131-3751	Sequence 3751, Ap

ALIGNMENTS

```
RESULT 1
US-08-961-083-1
; Sequence 1, Application US/08961083
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OR INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961, 083
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-961-083-1

Query Match 100.0%; Score 1999; DB 4; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Gaps 0;
```

|||||
Db 1 TAAATCTAGACAAATAAAATCAACATCGCTGAGCTTGGGTTCTGAAGCGCGCCGCA 60
QY 61 TGCCCAAGTAATGATATTCACAGATTTGGTTAAAGCAATGGTTCTATTCGAAGACA 120
|||||
Db 61 TGCCCAAGTAATGATATTCACAGATTTGGTTAAAGCAATGGTTCTATTCGAAGACA 120
QY 121 TGCGCTCTTGACACAGGGGGATTGATACATCCGTATCCTGGAGCTTCTTGCGCAA 180
121 TGCGCTCTTGACACAGGGGGATTGATACATCCGTATCCTGGAGCTTCTTGCGCAA 180
Db 121 TGCGCTCTTGACACAGGGGGATTGATACATCCGTATCCTGGAGCTTCTTGCGCAA 180
QY 181 TGCGCAAGCAATTCCTCCAGGTGATGATACCTTCACCCACAGTGTGATTAAGTGAC 240
181 TGCGCAAGCAATTCCTCCAGGTGATGATACCTTCACCCACAGTGTGATTAAGTGAC 240
Db 181 TGCGCAAGCAATTCCTCCAGGTGATGATACCTTCACCCACAGTGTGATTAAGTGAC 240
QY 241 TTAATTTTCACTTCGACTTCGACAGACTATTTCTCGTAAGGCTGAGAACCTTGTT 300
241 TTAATTTTCACTTCGACTTCGACAGACTATTTCTCGTAAGGCTGAGAACCTTGTT 300
Db 241 TTAATTTTCACTTCGACTTCGACAGACTATTTCTCGTAAGGCTGAGAACCTTGTT 300
QY 301 AGCGATTCACTTGAACAAAGCAACAGCAAGAAATCTTGACTATATATAA 360
301 AGCGATTCACTTGAACAAAGCAACAGCAAGAAATCTTGACTATATATAA 360
Db 301 AGCGATTCACTTGAACAAAGCAACAGCAAGAAATCTTGACTATATATAA 360
QY 361 GGCTCTACATCTTATGAGGAACTATGGAATGCAAGACAGCTCAAAACTACTATGTTAA 420
361 GGCTCTACATCTTATGAGGAACTATGGAATGCAAGACAGCTCAAAACTACTATGTTAA 420
Db 361 GGCTCTACATCTTATGAGGAACTATGGAATGCAAGACAGCTCAAAACTACTATGTTAA 420
QY 421 AGACCTCAATATTTAAGTTTACCTGAGTTCAGCTTGGCTGGAGTTCGCTGAGGACC 480
421 AGACCTCAATATTTAAGTTTACCTGAGTTCAGCTTGGCTGGAGTTCGCTGAGGACC 480
Db 421 AGACCTCAATATTTAAGTTTACCTGAGTTCAGCTTGGCTGGAGTTCGCTGAGGACC 480
QY 481 AAACCAATATGACCCCTATTCACATCCAGAAGACGCCAAGCCGCAAACTTGCTT 540
481 AAACCAATATGACCCCTATTCACATCCAGAAGACGCCAAGCCGCAAACTTGCTT 540
Db 481 AAACCAATATGACCCCTATTCACATCCAGAAGACGCCAAGCCGCAAACTTGCTT 540
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGTGAACAGATATGAGAAACACTCAATAC 600
541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGTGAACAGATATGAGAAACACTCAATAC 600
Db 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGTGAACAGATATGAGAAACACTCAATAC 600
QY 601 ACCAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 660
601 ACCAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 660
Db 601 ACCAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 660
QY 661 TAAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 720
661 TAAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 720
Db 661 TAAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 720
QY 721 AACTGGGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 780
721 AACTGGGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 780
Db 721 AACTGGGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 780
QY 781 TTAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 840
781 TTAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 840
Db 781 TTAATTAAGTATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 840
QY 841 TGTGATGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 900
841 TGTGATGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 900
Db 841 TGTGATGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 900
QY 901 TGTGATGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 960
901 TGTGATGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 960
Db 901 TGTGATGATGATGATGATCAAAAGTCTCAATAGCAAGTAATTCCTCTTACATGTA 960
QY 961 ACCGATACAGACTATGCTGCTGCTGAGTACGAGTGTGATGATGATGATGATGATGAT 1020
961 ACCGATACAGACTATGCTGCTGCTGAGTACGAGTGTGATGATGATGATGATGATGAT 1020
Db 961 ACCGATACAGACTATGCTGCTGCTGAGTACGAGTGTGATGATGATGATGATGATGAT 1020
QY 1021 CGTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
1021 CGTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 CGTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 GGGCTACTTTGGCAATCATCCTTGCAATACGCTTGCAAACTGCGAAACGCTCCAGC 1140
1081 GGGCTACTTTGGCAATCATCCTTGCAATACGCTTGCAAACTGCGAAACGCTCCAGC 1140

Db 1081 GGGCTACTTTGGCAATCATCCTTGCAATACGCTTGCAAACTGCGAAACGCTCCAGC 1140
QY 1141 CGTGAAGAACTATAACAGAGTCGAGCTCAACCCGCGCAAGACTTCTTAATAGTCTAG 1200
1141 CGTGAAGAACTATAACAGAGTCGAGCTCAACCCGCGCAAGACTTCTTAATAGTCTAG 1200
Db 1201 AATGACTACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1201 AATGACTACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1201 AATGACTACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CAAAAAATATGAGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1261 CAAAAAATATGAGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1261 CAAAAAATATGAGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 TGAAGCTTACTATTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1321 TGAAGCTTACTATTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1321 TGAAGCTTACTATTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 AGAGTCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
1381 AGAGTCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Db 1381 AGAGTCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
QY 1441 CATGATGAAGAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
1441 CATGATGAAGAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Db 1441 CATGATGAAGAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 1501 TCAGGCTGATGAAGAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1501 TCAGGCTGATGAAGAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 TCAGGCTGATGAAGAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 GACCTCTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
1561 GACCTCTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db 1561 GACCTCTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 GGCTGATGAGCAAGGCTATTTCAACCGTCTGACACCACTGATGATGATGATGATGATGAT 1680
1621 GGCTGATGAGCAAGGCTATTTCAACCGTCTGACACCACTGATGATGATGATGATGATGAT 1680
Db 1621 GGCTGATGAGCAAGGCTATTTCAACCGTCTGACACCACTGATGATGATGATGATGATGAT 1680
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
1681 CGCTGCCAAAGTTTACCGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
Db 1681 CGCTGCCAAAGTTTACCGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1800
1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1800
Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1800
QY 1801 TTCTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
1801 TTCTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 1801 TTCTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
QY 1861 ATCAGATAGTTCATCACTTCAGAGTCTAGCTTCACACCTCCAGACCAATATATGATGAC 1920
1861 ATCAGATAGTTCATCACTTCAGAGTCTAGCTTCACACCTCCAGACCAATATATGATGAC 1920
Db 1861 ATCAGATAGTTCATCACTTCAGAGTCTAGCTTCACACCTCCAGACCAATATATGATGAC 1920
QY 1921 TACCAATCTCAACCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
1921 TACCAATCTCAACCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 1921 TACCAATCTCAACCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
QY 1981 TCCTCAACGACGACCA 1999
1981 TCCTCAACGACGACCA 1999
Db 1981 TCCTCAACGACGACCA 1999

RESULT 2

US-10-158-844-145/c
; Sequence 145, Application US/10158844

GENERAL INFORMATION:

; APPLICANT: Kunsch et al.

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESSES:

; ADDRESS: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue


```

Db 8147 GCGTGTATGACAGGCTATTCATACCGCTGACACACCTTAGGCAATGGCCTTACGGT 8088
QY 1681 CCGTCGCAAAAGTTTACCGCTGTATGATGACCTAGCTGTGAGGAAGCAATCCAGAGA 1740
Db 8087 CCGTCGCAAAAGTTTACCGCTGTATGATGACCTAGCTGTGAGGAAGCAATCCAGAGA 8028
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCATTTAAAAATGCTGTCG 1800
Db 8027 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCATTTAAAAATGCTGTCG 7968
QY 1801 TTCTAGTGTGAGACTACCTGCTCCACACACCCCATCACTGAAAGTTCAAGCTATC 1860
Db 7967 TTCTAGTGTGAGACTACCTGCTCCACACACCCCATCACTGAAAGTTCAAGCTATC 7908
QY 1861 ATCAGATAGTTCACTTCAGCTAGCTAGCTCAGCACTCCAGACACAATATAGTAGCAG 1920
Db 7907 ATCAGATAGTTCACTTCAGCTAGCTAGCTCAGCACTCCAGACACAATATAGTAGCAG 7848
QY 1921 TACCAATCCTAACAATATATACGCAACATCAATACACCCCTGATCAACAATAATGNA 1980
Db 7847 TACCAATCCTAACAATATATACGCAACATCAATACACCCCTGATCAACAATAATGNA 7788
QY 1981 TCCTCAACGACGACACCA 1999
Db 7787 TCCTCAACGACGACACCA 7769

```

RESULT 3

US-10-121-120-27/c

Sequence 27, Application US/10121120

GENERAL INFORMATION:

APPLICANT: Bergeron, Michel G.

APPLICANT: Ouellette, Marc

APPLICANT: Roy, Paul H.

TITLE OF INVENTION: Specific and Universal Probes and Amplification

TITLE OF INVENTION: Primers

TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial

TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical

TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro

FILE REFERENCE: 12287_31

CURRENT APPLICATION NUMBER: US/10/121,120

CURRENT FILING DATE: 2002-04-11

PRIOR APPLICATION NUMBER: 09/452,599

PRIOR FILING DATE: 1999-12-01

PRIOR APPLICATION NUMBER: 08/304,732

PRIOR FILING DATE: 1994-09-12

NUMBER OF SEQ ID NOS: 177

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 27

LENGTH: 9100

TYPE: DNA

ORGANISM: Haemophilus influenzae

US-10-121-120-27

Query Match 5.9%; Score 117.2; DB 7; Length 9100;
 Best Local Similarity 50.3%; Pred. No. 2.6e-22;
 Matches 349; Conservative 0; Mismatches 333; Indels 12; Gaps 2;

```

QY 22 TCACTCATTTGCTGACTTGGTTCTGCAAGCCGCTCAATGCCCAAGCTAATGATATTC 81
Db 3498 TAAATTAATTTGGCGAAGTGGGTGAGCAAGCGCTATTCAGTAAGAAATTAAGCGATGTGC 3439
QY 82 CACACATTTGGTTAAGGCAATGTTCTATGAGAGCAATGCTCTTCTGAGACAGAGG 141
Db 3438 ACAAGCTTAATTTAGCGCTTTTTCAGCGAGAGACAGTCTTTTTCAGATCATCAGG 3379
QY 142 GATTGATACCATCCGTATCTGAGAGCTTTCTTGGCAATCTGCAAA--GCAATTCCT 198
Db 3378 ATTACACCTATCGGCAATTTGCCGATGTTTGTGCGAGAGTAAATGGCGGTGCATC 3319
QY 199 CCAAGGTGATCACTCTCACCAACAGTTGATTAAGTTGACTTTTCAACTTCGAC 258

```

```

Db 3318 ACAAGCGCAAGTAGATTACTCAACAAATTAAGCGCTAATCTTTTCTTA-----AC 3268
QY 259 TTCCGACGACGATTTCTCTGTAAGGCTCAGAGAGCTGGTATAGGATTCAGTATGACA 318
Db 3267 CTCGAAATAACCATTTATTCGTAAAGCTGTGAGAGCGCTTGGGTGGAATTCAGAAA 3208
QY 319 AAAAGCAACAGCAAGAAATCTTGACCTACTATATATAATTAAGTGTCTACATGTCTAATG 378
Db 3207 TACTCTCAACAAACAAGAAATATTAGCTTTATTTAAACAAATCTTTTAGGCGATCG 3148
QY 379 GAACATATGAAATGACAGACAGCTCAAACTACTATGTGTAAAGACCTCAATATTTAAG 438
Db 3147 TTCTATGTGTTGACAGCGGACACAAACCTATTTGCTTAATCATTTGAATGAATGAC 3088
QY 439 TTACCTCGATTAGCTTGGCTGAGATGCTGAGACCAACCAATATAGCCCTTA 498
Db 3087 CTATCGGAATAGCGATTTATGCTGTGTTTACTTAAGACCTTCAACATAGACCCGCT 3028
QY 499 TTCACATCCAGAGAGCCCAAGACCGCGAAACTTGTCTTATCTGAATGAAATAATCA 558
Db 3027 TTATCTTAAACAGTTACGAAAGAGCGCGCAATGTGTGTAAAGCGGTATGTTAGTGA 2968
QY 559 AGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATACCAATTAATGATGACT 618
Db 2967 AAAATACATCACCAAGAAAGAAATATGATGCTGATGAAGAGCGATGTGGCGACTA 2908
QY 619 ACAAGTCTCAATACAGAGTAATTAACCTGCTTACATGATGATTAATTAACCTCAGAGAGT 678
Db 2907 TCACGCGCAAAATTTGAATTTTCGAGCCGATATGCTGATGAAGTGTGCTCAAGAAAT 2848
QY 679 CATCACTCAAGTTGAAGAAAGAGGCTATTAAC 712
Db 2847 GGTGGTGTCTTTGGGGAAGAAATGCTTACAC 2814

```

RESULT 4

US-09-540-209B-3719

Sequence 3719, Application US/09540209B

GENERAL INFORMATION:

APPLICANT: Gary L. Breton

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES

TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.1001-001

CURRENT APPLICATION NUMBER: US/09/540,209B

CURRENT FILING DATE: 2000-04-04

NUMBER OF SEQ ID NOS: 10444

SEQ ID NO 3719

LENGTH: 2370

TYPE: DNA

ORGANISM: B.fragilis

US-09-540-209B-3719

Query Match 3.0%; Score 59.2; DB 5; Length 2370;
 Best Local Similarity 46.9%; Pred. No. 3.8e-06;
 Matches 261; Conservative 0; Mismatches 283; Indels 12; Gaps 2;

```

QY 42 GTTTCGAAGCGCGGCTCAATGCCCAAGCTAATGATATTCACAGATTGGTTAAGGCAA 101
Db 233 gtaagaaatctggtataacaccttataatgaacttaccacaacatctgcaatgcat 292
QY 102 TCGTTTATCGAAGACATCGCTTCTTGACACAGAGGGAATGATACATCCGTATCC 161
Db 293 tgatgccaaggaagagcttcttacgaacatctgggtatcgatgcgaagaagcgctga 352
QY 162 TGGAGCTTTTCTT-----GCGCAATCTGCAGAAAGCAATTCCTCCAGAGTGATCACTC 215
Db 353 tagcgctgctggttaagcggtgcatgcatcagaataatcaggtggaagcagtlacac 412
QY 216 TCACCAACAGTTGATTA-----GTTGACTTACTTTTCAACTCTGCACTTCGACGCA 269
Db 413 ttccacaacagctgcgaagcaatgttaccgagcaagtlgtccagaaataacatgctgcagc 472

```



```

; SEQ ID NO 2857
;
; LENGTH: 1350
;
; TYPE: DNA
;
; ORGANISM: Candida albicans
;
US-10-179-131-2857

```

Query Match	2.2%;	Score	44.2;	DB	7;	Length	1350;
Best Local Similarity	60.3%;	Pred. No.	0.052;	Mismatches	48;	Indels	0;
Matches	73;	Conservative	0;	Mismatches	48;	Indels	0;

[illegible]

QY	1966	T	1966
Db	1106	t	1106

RESULT 10
US-09-673-476-438
; Sequence 438, Application US/09673476
; GENERAL INFORMATION:

: APPLICANT: BOCHRIESEN-BROSCH, ROLAND
 : APPLICANT: GORDON, STEPHEN
 : TITLE OF INVENTION: BILLIULT, ALAIN
 : TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
 : TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
 : TITLE OF INVENTION: DNA LIBRARY, APPLICATION TO THE DETECTION OF
 : TITLE OF INVENTION: MYCOBACTERIA.

CURRENT APPLICATION NUMBER: 09/673,476
 CURRENT FILING DATE: 2003-03-29
 PRIOR APPLICATION NUMBER: PCT/IB99/00740
 PRIOR FILING DATE: 1999-04-16
 PRIOR APPLICATION NUMBER: 09/060,756
 PRIOR FILING DATE: 1998-04-16
 NUMBER OF SEQ ID NOS: 743
 SOFTWARE: PatentIn Ver. 2.2
 SEQ ID NO: 438

```

? ORGANISM: Mycobacterium tuberculosis
? FEATURE:
? NAME/KEY: modified_base
? LOCATION: (136)
? OTHER INFORMATION: a, t, c or g
? FEATURE:
? NAME/KEY: modified_base
? LOCATION: (158)
? OTHER INFORMATION: a, t, c or g
? FEATURE:
? NAME/KEY: modified_base
? LOCATION: (203)
? OTHER INFORMATION: a, t, c or g
? OS-09-673-476-438

```

Query March	2.0%;	Score 40.6;	DB 5;	Length 223;
Best Local Similarity	48.6%;			
Matches 106; Conservative	0;	Pred. No. 0.28;		
		Mismatches 112;	Indels 0;	Gaps 0;

QY 281 AAGCTCAGGAACTTGGTTAGACATTCATTAGAACAAAAAGCAACCACAGAATAATC 340
||| || | | | | | | | | | | | |
Db 6 aagcttcgcgagatcggatggaactcaacgtcgacaagaaccttcaaaaaattcgaaatc 65

QY 341 TTGACTACTATTATTAATTAAGGTGTACATGTCTAATGTGGAACTTAGAATGCAGACAGCA 400

Db 66 ctgaacccgatactcgtgaaccggtcgtcgttcggcacaataccctcgttcgycgtgcaggaacgcg 125

OY 401 GCTCAAAACTACTATGTGTAAGACCTCAATATATTTAAAGTTTACCTGATTAGCCCTTGCTG 460

Db 126 gcgcacaacgctcctcgcgcatcaacgcgtccgcgancctgattcgcagcaacgcgcgcgtctg 185

OY 461 GCTGGAATGCTCTCAGGCGACCAAAACCATATATGACCCCTTA 498

Db 186 gcgcgcatagtgtgcatactnaccacgaacgcgtcaaacccgcta 223

RESULT 11
US-10-179-131-3015
; Sequence 3015, Application US/10179131
; Category: INNOVATION.

1. TITLE OF INVENTION: CANDIDA ALICANS NUCLEIC ACIDS AND POLYPEPTIDES.
 2. APPLICANT: GREENE, JONATHAN R.
 3. APPLICANT: ZENG, QIANDONG
 4. APPLICANT: NOLLING, JORR
 5. APPLICANT: KESSLER, MARCO
 6. APPLICANT: SHIMER JR., GEORGE H.
 7. APPLICANT: SHAW, KAREN J.
 8. ATTORNEY: HART, ROBERT A. S.

;;
;; CLIB NUMBER: 270-9-400
;; CURRENT APPLICATION NUMBER: US/10/179,131
;; CURRENT FILING DATE: 2002-06-21
;; NUMBER OF SEQ ID NOS: 10194
;; SEQ ID NO 3015
;; LENGTH: 1611
;; TYPE: DNA
;; ORGANISM: Candida albicans
;; US-10-179-131-3015

Query Match	2.0%;	Score 40.4;	DB 7;	Length 1611;
Best Local Similarity	52.4%;	Pred. No. 0.65;		
Matches 89; Conservative	0;	Mismatches 81;	Indels 0;	Gaps 0;

OY	1830	AACCCCATCACTGAAAGTTTCAGGCTCATCAATCAATATGTTTCAACTTTCACAGTCTAGCT	1889
	152	atccaccaccacacacacacaccccttcattcacaacaaatgatttaccctcatgtgccaccaa	211
Db	1890	CAACCATCTCCAGACGACAAATAATATAGTAGCAGCTACCAATCTATAAATAATATAGCAACAAT	1944
OY	212	tactctgcacatccttaccctcctatatacaataatataatcgaatgtgtaacttaactgaat	271
Db	1950	CAAAATCAACCCCTGATCAACAAATAATAGATCTCTCAACACAGACAAACCA	1999
OY	272	tgaataacaacaatcagcaacaacaacaacaatcgtgtaacatggaacgcga	321

```

RESULT 12
US-10-027-632-5967/c
: sequence 5967, Application US/10027632
:
: GENERAL INFORMATION:
:
: APPLICANT: Wang, David G.
:
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: POLYMORPHISMS IN THE HUMAN GENOME
:
: TITLE OF INVENTION: Polymorphisms In the Human Genome
:
: FILE REFERENCE: 108827.129
:
: CURRENT APPLICATION NUMBER: US/10/027.632

```

PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/198,676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193,483
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: US 60/185,218
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: US 60/167,363
 PRIOR FILING DATE: 1999-11-23

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 19:00:18 ; Search time 2127.82 Seconds
(without alignments)
12679.838 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTACGACATATAA.....ATCCTCACGACACACCA 1999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59.4	3.0	277	10	BM398655 5009-0-48
2	54.4	2.7	939	12	CNS00CNG
3	53.6	2.7	529	12	BM398931 AG-ND-121
4	52.8	2.6	447	9	AM502461 UI-HF-BRO
5	52.6	2.6	630	12	BM393736 AG-ND-136
6	51.8	2.6	691	12	BM384275 AG-ND-145
7	49.4	2.5	922	12	CNS0073M
8	49.2	2.5	1101	12	CNS00FXE
9	47.2	2.4	612	9	BE195101
10	47	2.4	934	3	BI644440
11	47	2.4	1749	10	BM415379
12	46.8	2.3	860	12	CNS018FL
13	46.6	2.3	674	12	BM399409
14	46.6	2.3	884	12	CNS0060U
15	46.4	2.3	751	12	BM381742
16	45.6	2.3	485	12	BM242511
17	44.8	2.2	628	12	BM377839

18	44.4	2.2	419	9	AU033391	AU033391
19	44.4	2.2	544	9	AU037837	AU037837
20	44.4	2.2	895	12	CNS0071A	AL066286 Drosophila
21	43.8	2.2	988	12	CNS0072R	AL066743 Drosophila
22	43.6	2.2	475	10	BI594837	BI594837 Ag-LyC-44
23	43.4	2.2	472	10	BI114575	BI114575 602286235
24	43.4	2.2	582	9	AU060696	AU060696
25	43	2.2	572	12	A2042338	A2042338 RPI-23-3
26	43	2.2	1101	12	CNS00LOO	AL066807 Drosophila
27	42.8	2.1	1043	12	AG135568	AG135568 Pan trogl
28	42.4	2.1	1101	12	CNS0100X	AL098379 Drosophila
29	42.2	2.1	532	12	A2556585	A2556585 RPI-23-2
30	42.2	2.1	544	12	A2800772	A2800772 2M059A03
31	42	2.1	606	12	BM396411	BM396411 AG-ND-139
32	42	2.1	998	12	CNS0144M	AL103696 Drosophila
33	41.4	2.1	579	10	C25758	C25758 C25758 Dict
34	41.2	2.1	1101	12	CNS0160F	AL107217 Drosophila
35	40.8	2.0	414	9	A1590574	A1590574 tw13h08.x
36	40.8	2.0	1072	12	CNS0070B	AL067418 Drosophila
37	40.8	2.0	1166	12	AG126229	AG126229 Pan trogl
38	40.6	2.0	690	12	B07870	B07870 XK37T7 PB1u
39	40.6	2.0	829	9	AL569367	AL569367 AL569367
40	40.6	2.0	878	12	CNS0187R	AL108993 Drosophila
41	40.6	2.0	928	12	A2667900	A2667900 ENTG234TR
42	40.6	2.0	1101	12	CNS00LJT	AL068307 Drosophila
43	40.4	2.0	1036	12	CNS01ZIF	AL174336 Tetradon
44	40.4	2.0	1989	10	BF347552	BF347552 602020402
45	40.2	2.0	508	9	AU061871	AU061871 AU061871

ALIGNMENTS

RESULT 1
BM398655/c 277 bp mRNA linear EST 17-JAN-2002
LOCUS 5009-0-48-D02.t.1 Chilcoat/Turkewiltz cDNA (large fraction)
DEFINITION Tetrahymena thermophila cDNA, mRNA sequence.
ACCESSION BM398655.1 GI:18196708
VERSION EST.
KEYWORDS Tetrahymena thermophila.
SOURCE Tetrahymena thermophila
ORGANISM Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Tetrahymenina; Tetrahymena.

REFERENCE
AUTHORS I (bases 1 to 277)
Turkewiltz,A.P., Karrer,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel
,J. and Klobutcher,L.
EST from Tetrahymena thermophila, strain CU428.1, growing cells
unpublished (2002)
TITLE Contact: Turkewiltz AP
JOURNAL Molecular Genetics and Cell Biology
COMMENT University of Chicago
920 E. 58th Street, Chicago, IL 60637, USA
Tel: 773 702 4374
Fax: 773 702 3172
Email: apturkew@midway.uchicago.edu
Seq primer: 73.

FEATURES

source
1..277
Location/Qualifiers
/organism="Tetrahymena thermophila"
/strain="CU428.1"
/db_xref="taxon:5911"
/clone_lib="Chilcoat/Turkewiltz cDNA (large fraction)"
/note="Vector: Bluescript2 SK+, Details on library
preparation can be found in Chilcoat and Turkewiltz (2001)
Proc. Natl. Acad. Sci USA; 98: 8709-8713."
BASE COUNT 51 a 69 c 99 g 58 t
ORIGIN

Query Match 3.0%; Score 59.4; DB 10; Length 277;
Best Local Similarity 54.9%; Pred. No. 0.00025;

	Matches	117,	Conservative	0;	Mismatches	96;	Indels	0;	Gaps	0;
Qy	327	CCAACCAAGAAATCTT	GACTACTATATAAAT	AAGCTTACTATGCTATATGGAACTATG	386					
Db	257	CTAAACACAGATGCTT	GAGCGCTACCTCAATATATG	CGCTATTTGCTAGCGGTGCTACG	198					
Qy	387	GAATGCAGACAGACACT	CAAAACTACTATGTTAGTAAAGACCTCAATATATTAAGTTAACTTC	446						
Db	197	GGGTGAGATCTGACACT	CATCACTACTTTTGGGACCAACAGCGCCCAACTTGACCTGGCC	138						
Qy	447	AGTTAGCCTTGCTGGCT	GGATGATCCTCAAGGACCAACCAATATATGACCCCTATTTCATTC	506						
Db	137	AGTCGGCATGTGCTAC	TGGCTGGTACAGAAACCCGGTGGCTACAGACCCAGTAAACCAAC	78						
Qy	507	CAGAGCAAGCCCAAGAC	CGCCGCAACTGTGTC	539						
Db	77	CAGAGGCAGCCATGAAC	CGCCGCAATATATG	GAGCT	45					

RESULT	2
CNS00CNG	
LOCUS	939 bp DNA linear
DEFINITION	CNS00CNG Drosophila melanogaster genome survey sequence TE13 end of BAC # BACR2H16 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL059400
VERSION	AL059400.1
KEYWORDS	GI:4946964
SOURCE	GSS.
ORGANISM	fruit fly.
	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 939)

REFERENCE 1 (bases 1 to 939)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

COMMENT

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES      Location/Qualifiers
source        1. .939
```

BASE COUNT	ORIGIN	71 a	349 c	104 g	180 t	235 others
/organism="Drosophila melanogaster"						
/db_xref="taxon:7227"						
/clone_lib="RPCI-98"						
/clone="BACR26H16"						
/note"end : "TET3"						

Query Match	2.7%	Score	54.4	DB	12	Length	939
Best Local Similarity	10.4%	Pred. No.	0.0076				
Matches	21	Conservative	122	Mismatches	59	Indels	0
						Gaps	0

[illegible][illegible]

RESULT	3
BH398931/c	
LOCUS	BH398931
DEFINITION	AG-ND-12J05.TF ND-TAM Anopheles gambiae genomic clone AG-ND-12J05, DNA sequence.
FEATURES	529 bp DNA linear GSS 11-DEC-2001

ACCESSION	BH398931
VERSION	BH398931.1
KEYWORDS	GI:17345147
SOURCE	GSS.
ORGANISM	African malaria mosquito.
	Phelobotanella

ORGANISM
Anophelinae: gamuzae
Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
Pterygota: Neoptera: Endopterygota: Diptera: Nematocera: Culicoidae
; Anophelinae.
REFERENCE
1 (bases 1 to 529)

AUTHORS Shetty, J. , Malek, J. , Koo, H. , Collins, F. , Gardner, M. and Loftus, B. J.
 TITLE Direct Submission of BAC-end sequences from Anopheles gambiae
 JOURNAL Unpublished (2001)
 COMMENT Other_GSSS: AG-ND-121J5.TR

Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208
Fax: 301 838 3543

Email: bjo@fusenstifr.org
This clone is from an *A. gambiae* BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from *A. gambiae* PST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.
Seq primer: M13 For
Class: BAC ends.

FEATURES	location/qualifiers
source	1. .529

```

/organism="Anopheles gambiae"
/strain="PEST"
/db_xref="taxon:7165"
/clone="AG-ND-121J5"
/clone_1bp="ND-TAM"
/clone_vector="pECBAC1"
/site_1="HindIII"

BASE COUNT
ORIGIN
94 a      154 c      186 g      95 t

```

Query Match	2.7%	Score 53.6;	DB 12;	Length 529;
Best Local Similarity	53.9%;	Pred. No. 0.0096;		
Matches 110; Conservative	0;	Mismatches 94;	Indels 0;	Gaps 0;

QY 26 CTCATTGCTGACTTGGGTTCTGAAGCGCCGTCATATGCCCAAGCTATATGATATTCCACA 85
Db 278 CTCCTGGCGAATTCGGCGAGCGGCCCGCACCTGACACCCCATCAACGAGATCCCCAG 219

Qy 86 GATTGGTAAGCAATCGTTTCTATGGAAGCCATGGCTTTTCGACACAGGGGGATT 145
| | | | | | | | | | | | | | | |
Db 218 GTCATGACGATCGGTCTGCCATCAGGCACCCGCTTTTCGACACACGGTGCTG 159

OY 146 GATACCACGATTCCTGGAGCTTTCTTGGCAATCTGCAGAACTTCCTCAAGT 205
 DB 158 GACTACAGAGGCGATGCTGGCTGGCTGGCAACCTGGCGGCAAGATCAAGACCCAGGG 99
 OY 206 GATCACTCTCACCCACAGTTG 229
 DB 98 GCTCCACCATCACATGACGAGTG 75

RESULT 4
 AM502461 447 bp mRNA linear EST 01-MAR-2000
 LOCUS UI-HF-BR0P-ajv-b-08-0-UI.r1 NIH_MGC_52 Homo sapiens cDNA clone
 DEFINITION IMAGE:3075783 5', mRNA sequence.
 ACCESSION AM502461
 VERSION AM502461.1 GI:7116965
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 447)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Eco RI site shown at the beginning of the sequence.
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M.B. Soares Lab
 cDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbfp/image/image.html
 Seq primer: M13 Forward.

FEATURES

source

Location/Qualifiers

1..447

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3075783"

/clone_1ib="NIH_MGC_52"

/tissue_type="Lymph"

/cell_type="germinal center B cells"

/cell_line="MGC85"

/lab_host="DH10B (LTI)"

/note="Vector: p7773-Pac; Site_1: NotI; Site_2: Eco RI;
 Constructed from size fractionated cytoplasmic mRNA
 (7.4-9.5kb). Directionally cloned. Cells provided by
 Louis M. Staudt, Ph.D. Library preparation by Maria de
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT

73 a 134 c 156 g 83 t 1 others

ORIGIN

Query Match

2.6%; Score 52.8; DB 9; Length 447;

Best Local Similarity 53.1%; Pred. No. 0.014;

Matches 111; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

OY 329 AAGCAAGAATCTTGACTACTATATTAAGTGTACATGCTATGGAACTATGA 388
 DB 447 AAACGGAGATCTCTACCCGCTACCTGAACTGCTGCTGGGAAACAACACTGTCGGC 388
 OY 389 ATGCAGACAGACAGCTCAAACTACTATGTTAAAGACCTCAATATTTAAGTTTACCTCAG 448
 DB 387 ATCCAGAGAGCTGCGGAGCTACTTCTGGGGTCAACGATCGACTTGAACCTGGACGACG 328
 OY 449 TTAGCTTCTGCTGGTGAATGCTCAGGACCAACCAATATACCCCTTATTCATCCA 508
 DB 327 GGGGCTTCTGCGCGGAGTGTGAGTCGACGCGCGCTCAACCCNTACACCAACCC 268
 OY 509 GAAGCAGCCCAAGACCGCGAACTGTGT 537

DB 267 GAGGGCGCCTGCGCGGCGCAACTGTG 239

RESULT 5

BH393736

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 630)

Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.

Direct Submission of BAC-end sequences from Anopheles gambiae

Unpublished (2001)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjloftus@tigr.org

This clone is from an A. gambiae BAC library (ND-TAM) provided by

F.H. Collins and sequenced by The Institute for Genomic Research

(TIGR). The BAC library was generated from A. gambiae PEST strain

DNA. All DNA was extracted from newly hatched first instar larvae

to minimize the inclusion of DNA from microorganisms that inhabit

the gut. The DNA is derived from mixed sexes of larvae. The BAC

library was constructed at Texas A&M University BAC Center

University, College Station, Texas 77843-2123, USA using a HindIII

partial digest.

Seq primer: M13 For

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..630

/organism="Anopheles gambiae"

/strain="PEST"

/db_xref="taxon:7165"

/clone="AG-ND-13606"

/clone_1ib="ND-TAM"

/note="Vector: pECBAC1; Site_1: HindIII"

BASE COUNT

120 a 200 c 189 g 121 t

ORIGIN

Query Match

2.6%; Score 52.6; DB 12; Length 630;

Best Local Similarity 64.2%; Pred. No. 0.018;

Matches 79; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

OY 75 ATATCCACAGATTTGGTAAAGCAATGTTCTATCGAAGCAATCGTCTTGAC 134
 DB 504 AGATTTGCGCCCGCTGTGTAAAGCACTGTGGCAACGAGATCACGGTTCTAGAGC 563
 OY 135 ACAGGGGATTTGATACATCCGTATCTCGGAGCTTCTTGGCAATCTGCAAGAATT 194
 DB 564 ACCAGCGATGACATTCACCGGACCGCGTGGGTCTGTCATCATAGCTGTGGCGCAATT 623
 OY 195 CCC 197
 DB 624 CCC 626

RESULT 6
 BH384275 691 bp DNA linear GSS 10-DEC-2001
 LOCUS AG-ND-145F13.TR ND-TAM Anopheles gambiae genomic clone AG-ND-145F13
 DEFINITION , DNA sequence.
 ACCESSION BH384275

```

VERSION          BH384275.1   GI:17330417
KEYWORDS         GSS.
SOURCE           African malaria mosquito.
ORGANISM         Anopheles gambiae
                  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
                  Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides
                  ; Anophelinae.
REFERENCE        1 (bases 1 to 691)
AUTHORS          Shelby,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J
TITLE            Direct Submission of BAC-end sequences from Anopheles gambiae
JOURNAL          Unpublished (2001)
COMMENT          Other GSSs: AG-ND-145F13.FP
                  Contact: Brendan J Loftus
                  Department of Eukaryotic Genomics
                  The Institute for Genomic Research
                  9712 Medical Center Dr., Rockville, MD 20850, USA
                  Tel: 301 838 0208
                  Fax: 301 838 3543
                  Email: bjloftus@tigr.org
                  This clone is from an A. gambiae BAC library (ND-TAM) provided by
                  F.H. Collins and sequenced by The Institute for Genomic Research
                  (TIGR). The BAC library was generated from A. gambiae PEST strain
                  DNA. All DNA was extracted from newly hatched first instar larvae
                  to minimize the inclusion of DNA from microorganisms that inhabit
                  the gut. The DNA is derived from mixed sexes of larvae. The BAC
                  library was constructed at Texas A&M University BAC Center
                  University, College Station, Texas 77843-2123, USA using a HindIII
                  partial digest.
                  Seq primer: M13 Rev
                  Class: BAC ends.

FEATURES
    source                Location/Qualifiers
                            1..691
                            /organism="Anopheles gambiae"
                            /strain="PEST"
                            /db_xref="taxon:7165"
                            /clone="AG-ND-145F13"
                            /clone_id="ND-TAM"
                            /note="Vector: pECBAC1; Site_1: HindIII"

BASE COUNT      138 a       175 c       234 g       144 t

ORIGIN
Query Match          2.6%; Score 51.8; DB 12; Length 691;
Best Local Similarity 52.0%; Fred. No. 0.03;
Matches 116; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

OY      6 TCTAGACAATTAATACTCATTCGTCGACTTGCGGTCTTAGCGCCGCATAGCCCC 65
          ||||| | | | | | | | | | | | | | | | | | | | | | |
DB      293 TCTACTCCAGCAGAAGGCCCTCTGCTGGCGCAATTGGGAGAGAGCTCGCACCCGACGC 234
          || | | | | | | | | | | | | | | | | | | | | | |

OY      66 AAGCTAATGATTTCCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGCAATCGCT 125
          || | | | | | | | | | | | | | | | | | | | | | |
DB      233 CCATCAAGAGATCCCCCAAGGTCATGACGAGGAGCCGTCGTGSCATCGAAGACCCGCT 174
          || | | | | | | | | | | | | | | | | | | | | | |

OY      126 TCTTGACCACAGGGGGGAGTTGATACCATCCGATATCCTGGGACCTTTCTTGGCAATCTGC 185
          ||||| | | | | | | | | | | | | | | | | | | | |
DB      173 TCTTGACCATGCGCGCGGAGTACAAAGGGGACTGCTGCGCTCGCGTGCACCACTGG 114
          ||||| | | | | | | | | | | | | | | | | | | | |

OY      186 AAAGCAATTCCCTCAAGGTGATGATCACTCTCACCACAAGTT 228
          || | | | | | | | | | | | | | | | | | | | | |
DB      113 GCAGGTCAAGAGCCAGGGCGCGCTGACCATTCACCATGCAAGT 71

RESULT          7
CONS0073W/c     922 bp      DNA      linear      GSS 03-JUN--1999

DEFINITION      Drosophila melanogaster genome survey sequence Tef3 end of BAC #
LOCUS           BA014D09 of RpCl-98 library from Drosophila melanogaster (fruit
                 fly), genomic survey sequence.
ACCESSION       AL066784
VERSION         AL066784.1   GI:4945247
KEYWORDS        GSS.
SOURCE          fruit fly.
```

[illegible]


```

XX 13-SEP-2001 (Rel. 69, Created)
DT 13-SEP-2001 (Rel. 69, last updated, Version 1)
XX
DE OP165 Mixed Stage EST's from Globodera pallida, the potato cyst nematode
DE Globodera pallida cDNA, mRNA sequence.
XX
KM EST.
XX
OS Globodera pallida
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
XX
RN 1
RP 1-934
RA Heer J., Sosinski B., Pokrzywa R.M., Werry A., Opperman C.;
RT "Mixed Stage EST's from Globodera pallida, the potato cyst nematode";
RL Unpublished.
XX
CC Contact: Opperman, C
CC Center for the Biology of Nematode Parasitism
CC NC State University; IACR-Rothamsted
CC Campus Box 7616; Raleigh, NC 27695, USA
CC Tel: 919.515.9500
CC Fax: 919.515.9500
CC Email: warthog@unity.ncsu.edu
XX No homology found. ; gtl1_2pcn_r_F10_pcn_r_079.abi.seq.screen.
XX
FH Key Location/Qualifiers
FH
FT 1. 934
FT /db_xref="taxon:36090"
FT /note="Vector: lambda GT11. This is a collaborative effort
FT between IACR-Rothamsted and North Carolina State
FT University. The library was constructed from mixed stage G.
FT pallida in lambda GT11 by Paul Burroughs,
FT IACR-Rothamsted."
FT /organism="Globodera pallida"
FT /clone_lib="Mixed Stage EST's from Globodera pallida, the
FT potato cyst nematode"
XX
SQ Sequence 934 BP; 513 A; 262 C; 27 G; 132 T; 0 other;

```

```

Query Match 2.4%; Score 47; DB 3; Length 934;
Best Local Similarity 54.3%; Pred. No. 0.55;
Matches 95; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

```

OY 1825 ACAACACCCCCCACTGAAAGTTCATGCTATCATGATAGTTCACTTCACAGTC 1884
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 475 ACAAAACACAAACACACACACACACACACACACACACACACACACACACCA 534
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1885 TAGCTCAACCCCTCCAGCAAAATATAGTACGACTCAATCTTACATTAATACCA 1944
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 535 CACACCCCCCAACACACACACACACACACACACACACACACACACACACAA 594
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1945 ACAATCAATCAACCCCTGATCAACAAATCAGATCTCAACACGACCAACCA 1999
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 595 AAAAAAAAAAAAAACCCCACTCAACACACACACACACACACACACACACAA 649
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 11
LOCUS BM415379 1749 bp mRNA linear EST 28-JAN-2002
DEFINITION nematode Globodera pallida cDNA, mRNA sequence.
ACCESSION BM415379
VERSION BM415379.1 GI:18381989
KEYWORDS EST.
SOURCE Globodera pallida.
ORGANISM Globodera pallida.
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.

```

```

REFERENCE 1 (bases 1 to 1749)
AUTHORS Heer,J., Sosinski,B., Pokrzywa,R.M., Werry,A. and Opperman,C.
TITLE Mixed Stage EST's from Globodera pallida, the potato cyst nematode
JOURNAL Unpublished (2001)
COMMENT Contact: Opperman, C
Center for the Biology of Nematode Parasitism
NC State University; IACR-Rothamsted
Campus Box 7616; Raleigh, NC 27695, USA
Tel: 919.515.6699
Fax: 919.515.9500
Email: warthog@unity.ncsu.edu
gtl1_2pcn_r_F10_pcn_r_079.abi.

```

```

FEATURES
source
1. 1749
/organism="Globodera pallida"
/db_xref="taxon:36090"
/clone_lib="Mixed Stage EST's from Globodera pallida, the
potato cyst nematode"
note="Vector: lambda GT11. This is a collaborative effort
between IACR-Rothamsted and North Carolina State
University. The library was constructed from mixed stage
G. pallida in lambda GT11 by Paul Burroughs,
IACR-Rothamsted."

```

```

BASE COUNT 957 a 479 c 53 g 260 t
ORIGIN

```

```

Query Match 2.4%; Score 47; DB 10; Length 1749;
Best Local Similarity 54.3%; Pred. No. 0.72;
Matches 95; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

```

OY 1825 ACAACACCCCCCACTGAAAGTTCATGCTATCATGATAGTTCACTTCACAGTC 1884
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 475 ACAAAACACAAACACACACACACACACACACACACACACACACACACCA 534
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1885 TAGCTCAACCCCTCCAGCAAAATATAGTACGACTCAATCTTACATTAATACCA 1944
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 535 CACACCCCCCAACACACACACACACACACACACACACACACACACACACAA 594
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1945 ACAATCAATCAACCCCTGATCAACAAATCAGATCTCAACACGACCAACCA 1999
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 595 AAAAAAAAAAAAAACCCCACTCAACACACACACACACACACACACACACAA 649
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 12
LOCUS CNS018FL 860 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN13D02 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL109275
VERSION AL109275.1 GI:5629579
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Eukaryota; Neoptera; Endopterygota; Diptera; Brachyera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 860)

```

```

REFERENCE 1
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre national de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billard at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
peloBAC11.

```

```

FEATURES
Location/Qualifiers

```

```

source                                1. .860
                                      /organism="Drosophila melanogaster"
                                      /plasmid="pbeloBAC11"
                                      /db_xref="taxon:7227"
                                      /clone_11b="DrosBAC"
                                      /clone="BACN13D02"
                                      /note="end : T7"
BASE COUNT      218 a      134 c      113 g      139 t      256 others
ORIGIN

Query Match                               2.3%; Score 46.8; DB 12; Length 860;
Best Local Similarity 28.1%; Pred. No. 0.6;
Matches 78; Conservative 68; Mismatches 132; Indels 0; Gaps 0;

QY 1722 AAGGAACAATCCAGAAGATTGGAATATACAGAGAGGGCTCTACAGAATAAGAGAAATTCG 1781
    :| | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 233 MAANAAMGAAAAMMAAAAAAAAAAAAAAACATTAATAAAGCCCCCGMSGAMMACAAMMMMA 292

QY 1782 TATTTAAAAATG TGTCGTCTTACGTGGAAGACTCACTGCCTCCACACAAACCCCATCAA 1841
    Db 293 AATCCGCCCMACCMAACCCACCAACAACCTTAGCGCAACGCCCCCGMAAMMMMCACMCMCMAAA 352

QY 1842 CTGAAGAAGTTCAAGCTCATCATCAGATAGTATGTTCAACTTCCAGCTAGCTCAACCACTCCA 1901
    :| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 353 NNANANANNAANNACNNNNNMNNNNMMNNMMNNMMAAAMAACAMMMMMCMCACCCMCMACHTGCC 412

QY 1902 GCACAATAATATAGTACGACTACCAATCTCTAAACAATATATAGCCACAATCAATCAACACC 1961
    :| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 413 ACNNAACTSMRTATMACCAACMMMMNCMAACMMMAAACAAMMAACMACMACMCSMAAMC 472

QY 1962 CTGATCAACAATAATCAGAAATCCTCAACCAGGACAACCA 1999
    : : : | | | | : : : : | | | | : | | | | : | | |
Db 473 IMCMMAACAACMACMCMCMCMSCMSCSACCCCAAMMAA 510

```

REFERENCE	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
<p> REFERENCES Anthonis TITLE JOURNAL COMMENT </p>	<p> BH399409 AG-ND-149L20 </p>	<p> 674 bp linear GSS 11-DEC-2001 AG-ND-149L20.TR ND-TAM Anopheles gambiae genomic clone AG-ND-149L20 , DNA sequence. </p>	<p> BH399409 BH399409.1 </p>	<p> GI:17345625 GSS. </p>	<p> African malaria mosquito. Anopheles gambiae Anopheles. 1 (bases 1 to 674) Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J. Direct Submission of BAC-end sequences from Anopheles gambiae Unpublished (2001) Other GSSs: AG-ND-149L20.TF </p>	<p> Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel.: 301 838 0208 Fax: 301 838 3543 Email: b.loftus@tigr.org </p>	<p> This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest. Seq primer: M13 Rev Class: BAC ends. </p>

[illegible]

RESULT	14
CNS00600/c	
LOCUS	
DEFINITION	CNS00600 Drosophila melanogaster genome survey sequence T7 end of BAC #
ACCESSION	BACRI4N21 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION	AT065923
KEYWORDS	GSS.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS	Pteriygota; Neoptera; Endopterygota; Diptera; Brachycera;
JOURNAL	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
TITLE	1 (bases 1 to 884)
COMMENT	Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosawa and Aron Mammosses in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the library pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES	
source	Location/Organism 1..884 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="RPCI-98" /clone="BACRI4N21" /note="end : T7"
BASE COUNT	230 a 62 c 139 g 124 t 329 others
ORIGIN	

Oy	61	TCGCCAAGCTAATATATATTCGCCACAGATTTGGTTAAGGCAATCTGTTCTATGGAAGACA	120
Db	61	TGCCCAAGCTAATATATATTCGCCACAGATTTGGTTAAGGCAATCTGTTCTATGGAAGACA	120
Oy	121	TCGCTTCATGCACACAGAGGGGAGATGTATACCATCCGATCTCGGAGGCTTCTTGCGCA	180
Db	121	TCGCTTCATGCACACAGAGGGGAGATGTATACCATCCGATCTCGGAGGCTTCTTGCGCA	180
Oy	181	TCGCAAAAGCAATTCCTCCAGAGGTGATCACTCCACCAACAGTTGATTAAGTTGAC	240
Db	181	TCGCAAAAGCAATTCCTCCAGAGGTGATCACTCCACCAACAGTTGATTAAGTTGAC	240
Oy	241	TTACTTTTCAACCTTGACCTCCAGCCAGACATATTTCTCGTAAGGCTCAGAAAGCTTGCT	300
Db	241	TTACTTTTCAACCTTGACCTCCAGCCAGACATATTTCTCGTAAGGCTCAGAAAGCTTGCT	300
Oy	301	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATTAATTA	360
Db	301	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATTAATTA	360
Oy	361	GGCTACATGTCATATGGAAGCTATGGAATGCAGACAGCAGCTCAAAACTACTATGTAA	420
Db	361	GGCTACATGTCATATGGAAGCTATGGAATGCAGACAGCAGCTCAAAACTACTATGTAA	420
Oy	421	AGACCTCAATATTTAAGTTTACTCAGTTAGCCTTGCTGGTGGAAATCCTCAGGACAC	480
Db	421	AGACCTCAATATTTAAGTTTACTCAGTTAGCCTTGCTGGTGGAAATCCTCAGGACAC	480
Oy	481	AAACCAATATGACCCCTTAATTCACATCCAGAGAGAGCCCAAGACCGCCGAATTTGGTCTT	540
Db	481	AAACCAATATGACCCCTTAATTCACATCCAGAGAGAGCCCAAGACCGCCGAATTTGGTCTT	540
Oy	541	ATCTGAAATGAAAAATCAGAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC	600
Db	541	ATCTGAAATGAAAAATCAGAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC	600
Oy	601	ACCAATTACATGATGACTACAAAGTCCAAATCAGAACATTAATTAATTAATTAATTAATTA	660
Db	601	ACCAATTACATGATGACTACAAAGTCCAAATCAGAACATTAATTAATTAATTAATTAATTA	660
Oy	661	TAATTAACCTCAAGGAAGTCTCATCATCAATCAAGTTGAGAAAGAACAGGCTATATACCTACAC	720
Db	661	TAATTAACCTCAAGGAAGTCTCATCATCAATCAAGTTGAGAAAGAACAGGCTATATACCTACAC	720
Oy	721	AACTGGATGATGCTTACACAAATGTAGACCAAGAAAGCTCAAAAACATCTGGGATAT	780
Db	721	AACTGGATGATGCTTACACAAATGTAGACCAAGAAAGCTCAAAAACATCTGGGATAT	780
Oy	781	TTCAATATACAGACGAATAAGCTTGGCTATCCAGACGATGGAATTTGCAAGTGGCTTACACAT	840
Db	781	TTCAATATACAGACGAATAAGCTTGGCTATCCAGACGATGGAATTTGCAAGTGGCTTACACAT	840
Oy	841	TGTTGATGTTTCTTAACGCTAAAGTCAATGGCCAGCTAGAGAGACGCACTACATCAAGTAA	900
Db	841	TGTTGATGTTTCTTAACGCTAAAGTCAATGGCCAGCTAGAGAGACGCACTACATCAAGTAA	900
Oy	901	TGTTTCTTGGGAATTAACCAAGCAGTAGAAACAAACGCGACTGGGGATCAACTATGAA	960
Db	901	TGTTTCTTGGGAATTAACCAAGCAGTAGAAACAAACGCGACTGGGGATCAACTATGAA	960
Oy	961	ACCGATACAGACATATCCTCGCTGGATCGGATCGGATGCTTCAGANTCAACGCTACATAT	1020
Db	961	ACCGATACAGACATATCCTCGCTGGATCGGATCGGATGCTTCAGANTCAACGCTACATAT	1020
Oy	1021	CGTTACAGATGAGCCCTATTAATTAACCTCGGAGCAAAATTAATCTCGTTTATTAATCGGATAG	1080
Db	1021	CGTTACAGATGAGCCCTATTAATTAACCTCGGAGCAAAATTAATCTCGTTTATTAATCGGATAG	1080
Oy	1081	GGGCTACTTTGGCAACATCACTTGGCAATACGCCCTGCAACAATCGCGAAAGCTCCAGC	1140
Db	1081	GGGCTACTTTGGCAACATCACTTGGCAATACGCCCTGCAACAATCGCGAAAGCTCCAGC	1140

QY	1141	CGGAAACCTCTAACAAGGTGGAGCTCAACCGCGCAAGACTTTCTTAATGGCTATGG	1200
Db	1141	CGTGGAAACTCTTAACAAGGTGGAGCTCAACCGCGCAAGACTTTCTTAATGGCTATGG	1200
QY	1201	AATGCAGTACCCAAAGTATTCACTACTCTCAATATGCGCATTTTCAAGTAACACACGGAATCAGA	1260
Db	1201	AATGCAGTACCCAAAGTATTCACTACTCTCAATATGCGCATTTTCAAGTAACACACGGAATCAGA	1260
QY	1261	CAAAAAATATGGAGCAAGTAGTGAAAAAGATGGGTGCTGCTTACGCTGGCTTTGCAATAGG	1320
Db	1261	CAAAAAATATGGAGCAAGTAGTGAAAAAGATGGGTGCTGCTTACGCTGGCTTTGCAATAGG	1320
QY	1321	TGGAACCTTACTATTAACCAATGTATATCCATAAGTCGCTTTAGTATGGAGTGAAAA	1380
Db	1321	TGGAACCTTACTATTAACCAATGTATATCCATAAGTCGCTTTAGTATGGAGTGAAAA	1380
QY	1381	AGAGTTCTTAATATGTCGGAACCTCGTGCCATGAAGAAGACAGCCCTATATGATAGCCGA	1440
Db	1381	AGAGTTCTTAATATGTCGGAACCTCGTGCCATGAAGAAGACAGCCCTATATGATAGCCGA	1440
QY	1441	CATGATGAAAAACATCTTGACTTATGAGTAATGGACCAATGGCTATCTATGGCTGGCTCC	1500
Db	1441	CATGATGAAAAACATCTTGACTTATGAGTAATGGACCAATGGCTATCTATGGCTGGCTCC	1500
QY	1501	TCAGGCTGCTAAAAACAGAACTCTCTAACTATACAGACGAGGAATTTGAAAAACCAATCAA	1560
Db	1501	TCAGGCTGCTAAAAACAGAACTCTCTAACTATACAGACGAGGAATTTGAAAAACCAATCAA	1560
QY	1561	GACCTCTCATATTTCTAGCACTGTATGTAAGTAATTTGCTGGCTTACCGCTAAATATTCAAT	1620
Db	1561	GACCTCTCATATTTCTAGCACTGTATGTAAGTAATTTGCTGGCTTACCGCTAAATATTCAAT	1620
QY	1621	GGCTGTATGGACAGGCTATTTCTTAACGGTGTGACACACTTTGAGCAATGGCTTACCGT	1680
Db	1621	GGCTGTATGGACAGGCTATTTCTTAACGGTGTGACACACTTTGAGCAATGGCTTACCGT	1680
QY	1681	CGGTGCCAAAGTTTACCGGCTCTATGATGACCTTACCTGTGTGAAGGACCAATCCGAGA	1740
Db	1681	CGGTGCCAAAGTTTACCGGCTCTATGATGACCTTACCTGTGTGAAGGACCAATCCGAGA	1740
QY	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTGGCG	1800
Db	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTGGCG	1800
QY	1801	TTTACAGTGGAACTCACCTCTCCACACAAACCCCATCACTAAGTGAAGTTCAAGCTCATC	1860
Db	1801	TTTACAGTGGAACTCACCTCTCCACACAAACCCCATCACTAAGTGAAGTTCAAGCTCATC	1860
QY	1861	ATCAGATATGTTCAACTTCACAGTCTTAGCTCAACCACTCCAAAGCAATAATATAGTAGAC	1920
Db	1861	ATCAGATATGTTCAACTTCACAGTCTTAGCTCAACCACTCCAAAGCAATAATATAGTAGAC	1920
QY	1921	TACCAATCTTACCAATTAATACGCAACAAATCAAAATTAACCCCTGTATCAACAAATACAGA	1980
Db	1921	TACCAATCTTACCAATTAATACGCAACAAATCAAAATTAACCCCTGTATCAACAAATACAGA	1980
QY	1981	TCCTTCAACGACGACACCA 1999	
Db	1981	TCCTTCAACGACGACACCA 1999	

RESULT 2
 US-08-461,435-5
 Sequence 5, Application US/08461435
 Patent No. 6027906
 GENERAL INFORMATION:
 APPLICANT: Balganesb, Tanjore S
 APPLICANT: Town, Christine
 TITLE OF INVENTION: No. 6027906el polypeptides
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: White & Case
 STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/481,435
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2049 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Streptococcus pneumoniae
STRAIN: PM 1
IMMEDIATE SOURCE:
LIBRARY: PCR cloning
CLONE: pARC 0512 Soluble PBP 1A del 38
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2049
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 1..2046
US-08-481-435-5

Query Match 99.1%; Score 1981.4; DB 3; Length 2049;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATAAAATCAATCTGCTGCTGGTTCTGAACGCCGCGTCAA 60
DB 48 TAAATCTAGACAAATAAAATCAATCTGCTGCTGGTTCTGAACGCCGCGTCAA 107
QY 61 TCCCAACTATGATATATCCACAGATTGGTTAAGCAATCGTTCTATGTAAGACCA 120
DB 108 TCCCAACTATGATATATCCACAGATTGGTTAAGCAATCGTTCTATGTAAGACCA 167
QY 121 TCGCTTCTGACACGAGGGGATGATACATCCGATCCGAGGCTTCTTCCGCA 180
DB 168 TCGCTTCTGACACGAGGGGATGATACATCCGATCCGAGGCTTCTTCCGCA 227
QY 181 TCTGCAAGCAATTCCTCCCAAGTGATCACTCCACCAACAGTTGATTAGTTGAC 240
DB 228 TCTGCAAGCAATTCCTCCCAAGTGATCACTCCACCAACAGTTGATTAGTTGAC 287
QY 241 TTACTTTCACTTCGACTTCGACACAGACTATTTCTGTAAAGGCTCAGGAAGCTTGGT 300
DB 288 TTACTTTCACTTCGACTTCGACACAGACTATTTCTGTAAAGGCTCAGGAAGCTTGGT 347
QY 301 AGGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATA 360

DB 348 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATA 407
QY 361 GGTCTACATGTTCTAATGGGAATATGGAATGACAGACAGCTCAAACTACTATGTTAA 420
DB 408 GGTCTACATGTTCTAATGGGAATATGGAATGACAGACAGCTCAAACTACTATGTTAA 467
QY 421 AGACCTCATATATTAAGTTACCTCAGTTAGCTTGGGCTGGGAATGCCAGGACAC 480
DB 468 AGACCTCATATATTAAGTTACCTCAGTTAGCTTGGGCTGGGAATGCCAGGACAC 527
QY 481 AAACCAATATGACCCCTATTCATCCAGAAAGCAGCCCAAGACCCGGAACCTTGGCTT 540
DB 528 AAACCAATATGACCCCTATTCATCCAGAAAGCAGCCCAAGACCCGGAACCTTGGCTT 587
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCAGTAAATAC 600
DB 588 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCAGTAAATAC 647
QY 601 ACCAATTCAGTGGAGCTACAAAGTCTCAAAATCAGCAAGTATTAACCTGCTACATGGA 660
DB 648 ACCAATTCAGTGGAGCTACAAAGTCTCAAAATCAGCAAGTATTAACCTGCTACATGGA 707
QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAACAGGCTATTAACCTACTCAC 720
DB 708 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAACAGGCTATTAACCTACTCAC 767
QY 721 AACTGGATGATGTCCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
DB 768 AACTGGATGATGTCCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 827
QY 781 TTACATACAGAGCAATACGCTGCTATCCAGACATGAATGAATGCAAGTCCCTACAT 840
DB 828 TTACATACAGAGCAATACGCTGCTATCCAGACATGAATGAATGCAAGTCCCTACAT 887
QY 841 TGTGATGTTTCTAAGCGGTAAGTCAATGCCAGTAGAGCAGCCCATCACTCAAGTAA 900
DB 888 TGTGATGTTTCTAAGCGGTAAGTCAATGCCAGTAGAGCAGCCCATCACTCAAGTAA 947
QY 901 TGTTCCTCTGGGAATTAACCAAGCAGTAGAAACAAACCGGCTGGGATCAATATGAA 960
DB 948 TGTTCCTCTGGGAATTAACCAAGCAGTAGAAACAAACCGGCTGGGATCAATATGAA 1007
QY 961 ACCGATCAAGCAATATGCTCTGCTGCTGGAGTACGCTGCTCAATCACTGCTACAT 1020
DB 1008 ACCGATCAAGCAATATGCTCTGCTGCTGGAGTACGCTGCTCAATCACTGCTACAT 1067
QY 1021 GCTTACAGTACAGCCCTATTAATCACTACCTGGGACAAATACCTGTTTAACTGGGATGA 1080
DB 1068 GCTTACAGTACAGCCCTATTAATCACTACCTGGGACAAATACCTGTTTAACTGGGATGA 1127
QY 1081 GGGCTACTTTGGCAACATCACTGCTCAATACGCCCTGCAACAAATGCGGAACGCTCCAC 1140
DB 1128 GGGCTACTTTGGCAACATCACTGCTCAATACGCCCTGCAACAAATGCGGAACGCTCCAC 1187
QY 1141 CGTGAACCTTAACCAAGGTGGGACTCAACCGGCGCAAGACTTTCCTAAATGGTCTAG 1200
DB 1188 CGTGAACCTTAACCAAGGTGGGACTCAACCGGCGCAAGACTTTCCTAAATGGTCTAG 1247
QY 1201 AATGACTACCAATTAATCACTACTCAATGCAATTCATCAATCAACAAACGAATACGA 1260
DB 1248 AATGACTACCAATTAATCACTACTCAATGCAATTCATCAATCAACAAACGAATACGA 1307
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGATGAGCTGCTGTTAGCGCTTTGCAAAATGG 1320
DB 1308 CAAAAAATATGAGCAAGTAGTGAAGATGAGCTGCTGTTAGCGCTTTGCAAAATGG 1367
QY 1321 TGGAACTTACTATTAACCAATGATATATCAATGAATGCTTATGATGGAGGAATA 1380
DB 1368 TGGAACTTACTATTAACCAATGATATATCAATGAATGCTTATGATGGAGGAATA 1427
QY 1381 AGAGTTCTATATGTCGGAACCTGTCATGAAAGAAAGCAAGGCTATATGATGACGA 1440

Db	1428	AGAGTTCCTAAATGTGGAACTGTGGCAATGAAAGAAAGACAGGCTAATATGATGACGA	1487
Oy	1441	CATGATGAAAACAGCTCTTGACTCTATGGAACGTGAGCAAAATGCCATCTTGTGGCTCCC	15000
Db	1488	CATGATGAAAACAGCTCTTGATGATATGGAACTGGACGCAAAATGCCCTATCTTGTGGCTCCC	15478
Oy	1501	TCGAGCTGGTAAACAGAAACGCTTAACATATACAGCGAGGAAATTTGAAACCACATCAA	15668
Db	1548	TCGAGCTGGTAAACAGAAACGCTTAACATATACAGCGAGGAAATTTGAAACCACATCAA	16077
Oy	1561	GACCTTCGAATTTGTATGACACCTGATGAAACATATTTGCTGCTATACGCGTAATATTCAT	16200
Db	1608	GACCTTCGAATTTGTATGACACCTGATGAAACATATTTGCTGCTATACGCGTAATATTCAT	16677
Oy	1621	GGCTGTATGAGCAGGCTATTTCTAACCGTCTGACACACACTTTGAGCAATGGCTTACGGT	16800
Db	1668	GGCTGTATGAGCAGGCTATTTCTAACCGTCTGACACACACTTTGAGCAATGGCTTACGGT	17277
Oy	1681	CGCTGCCAAAGTTTACCGGCTCTATGATGACCTTCCTGTGTGAAGGAGCAATCCGGAAGA	17400
Db	1728	CGCTGCCAAAGTTTACCGGCTCTATGATGACCTTCCTGTGTGAAGGAGCAATCCGGAAGA	17877
Oy	1741	TTTGAATATATACAGAGGGGCTCTACAGAAATGGAGAATTCGATTTTAAAAATGGTCTCG	18000
Db	1788	TTTGAATATATACAGAGGGGCTCTACAGAAATGGAGAATTCGATTTTAAAAATGGTCTCG	18477
Oy	1801	TTTCTACTGGGAACCTCACCTCTCTCCACAAACACCCCATCATCACTGAAAGTTCAAGCTCATC	18600
Db	1848	TTTCTACTGGGAACCTCACCTCTCTCCACAAACACCCCATCATCACTGAAAGTTCAAGCTCATC	19077
Oy	1861	ATCAGATAGTTCAACCTCACAGCTCTAGCTCAACCACTCCCAACACAAATTAATAGTACAC	19200
Db	1908	ATCAGATAGTTCAACCTCACAGCTCTAGCTCAACCACTCCCAACACAAATTAATAGTACAC	19677
Oy	1921	TACCAATCTTACCAATATATACGCAACATTCAAATACACCCCTGATCAACAAATTCAGAA	19800
Db	1968	TACCAATCTTACCAATATATACGCAACATTCAAATACACCCCTGATCAACAAATTCAGAA	20277
Oy	1981	TTCCCAACGACGACACCA 1999	
Db	2028	TTCTTAAACGACACCA 2046	

```

1  ATTORNEY/AGENT INFORMATION:
2  NAME: Jackson Esq., David A.
3  REGISTRATION NUMBER: 26,742
4  REFERENCE/DOCKET NUMBER: 600-1-069 CIP
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE: 201 487-5800
7  TELEFAX: 201 343-1684
8  TELEX: 133521
9  INFORMATION FOR SEQ. ID NO.: 3:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 960 base pairs
12 TYPE: nucleic acid
13 STRANDEDNESS: both
14 TOPOLOGY: unknown
15 MOLECULE TYPE: DNA (genomic)
16 HYPOTHEetical: NO
17 ANTI-SENSE: NO
18 ORIGINAL SOURCE:
19 ORGANISM: Streptococcus pneumoniae
20 STRAIN: R6
21 IMMEDIATE SOURCE:
22 CLONE: SPRU42
23 FEATURE:
24 NAME/KEY: CDS
25 LOCATION: 1..960
26 US-08-245-511-3

```

```

1      RESULT 3
2      ; Sequence 3, Application US/08245511
3      ; Patent No. 5928900
4      ; GENERAL INFORMATION:
5      ; APPLICANT: Mesure, H Robert
6      ; APPLICANT: Pearce, Barbara J
7      ; APPLICANT: Tuomanen, Elaine
8      ; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
9      ; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
10     ; NUMBER OF SEQUENCES: 58
11     ; CORRESPONDENCE ADDRESS:
12     ; ADDRESSEE: Klauber & Jackson
13     ; STREET: 411 Hackensack Avenue
14     ; CITY: Hackensack
15     ; STATE: New Jersey
16     ; COUNTRY: USA
17     ; ZIP: 07601
18     ; COMPUTER READABLE FORM:
19     ; MEDIUM TYPE: Floppy disk
20     ; COMPUTER: IBM PC compatible
21     ; OPERATING SYSTEM: PC-DOS/MS-DOS
22     ; SOFTWARE: Patentin Release #1.0, Version #1.25
23     ; CURRENT APPLICATION DATA:
24     ; APPLICATION NUMBER: US/08/245,511
25     ; FILING DATE: 18-MAY-1994
26     ; CLASSIFICATION: 424
27     ; PRIOR APPLICATION DATA:
28     ; APPLICATION NUMBER: US 08/116,541
29     ; FILING DATE: 01-SEP-1994

```

Query Match:		47.28:	Score 944.2:	DB 2:	Length 960:
Best Local Similarity		99.7%:	Pred. No. 1.9e-278:		
Matches	946:	Conservative	0:	Mismatches	3:
				Indels	0:
				Gaps	
QY	1	TAAATCTAGCAGCAATAAAAAATCAACTCATCTGCTGACTTGGGTTCTGAAGCGCGCTCAA	60		
Db	12	TAAATCTAGCAGCAATAAAAAATCAACTCATCTGCTGACTTGGGTTCTGAAGCGCGCTCAA	71		
QY	61	TGCCCAACTATGATATTGCCACAGATTGGTTAAGGCAATGGTTCTATCGAAGCCA	120		
Db	72	TGCCCAACTATGATATTGCCACAGATTGGTTAAGGCAATGGTTCTATCGAAGCCA	131		
QY	121	TGCGTTCTTCGACGACAGAGGGGATGTGATACATCCGATCTCGGAGCTTCTTCGCCAA	180		
Db	132	TGCGTTCTTCGACGACAGGGGATGTGATACATCCGATCTCGGAGCTTCTTCGCCAA	191		
QY	181	TCTGCAAGCAATTCCTCTCCAGGTGATCACTCTACCCACAGTTGATTAAGTTGAC	240		
Db	192	TCTGCAAGCAATTCCTCTCCAGGTGATCACTCTACCCACAGTTGATTAAGTTGAC	251		
QY	241	TTACTTTTCAACTTCGACTTCGACCTCCGACAGACTATTTCTGTAAGGCTCAGAAAGCTTGTT	300		
Db	252	TTACTTTTCAACTTCGACTTCGACCTCCGACAGACTATTTCTGTAAGGCTCAGAAAGCTTGTT	311		
QY	301	AGCGATTAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATTA	360		
Db	312	AGCGATTAGTTAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATTA	371		
QY	361	GGTTCATATGCTATATGGGAACTTTGGAATCAGACAGACCTCAAAACTACTATGGTAA	420		
Db	372	GGTTCATATGCTATATGGGAACTTTGGAATCAGACAGACCTCAAAACTACTATGGTAA	431		
QY	421	AGACCTCAATATTTAAATTTTACCTGATTAAGCCTCTGGCTGGAAATGCTCAGGCAAC	480		
Db	432	AGACCTCAATATTTAAATTTTACCTGATTAAGCCTCTGGCTGGAAATGCTCAGGCAAC	491		
QY	481	AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGCCGCGGAATCTTGCTT	540		
Db	492	AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGCCGCGGAATCTTGCTT	551		
QY	541	ATCTGAATGAAAAATCAAGGCTCATCTCTGCGAAGCAGATGAGAAACAGTCAATAC	600		
Db	552	ATCTGAATGAAAAATCAAGGCTCATCTCTGCGAAGCAGATGAGAAACAGTCAATAC	611		
QY	601	ACCAATTACTGATGACTACAAAGCTCTCAAAATCGACAATTAATTAACCTCTTACATGGA	660		

|||||
Db 612 ACCAATTACTGATGGGCTACAAAGTCTCAATACAGAACTAATTAATACCCGTGTACATGGA 671
OY 661 TAATTACCTCAAGGAAGATCATCATCACTTAAGAGAAAGAGCTATTAACCTACTGC 720
Db 672 TAATTACCTCAAGGAAGATCATCATCACTTAAGAGAAAGAGCTATTAACCTACTGC 731
OY 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
Db 732 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 791
OY 781 TTACATACAGACGAATACGTTGCCATACAGACGATGAATTCGAATGCTTCTACCAT 840
Db 792 TTACATACAGACGAATACGTTGCCATACAGACGATGAATTCGAATGCTTCTACCAT 851
OY 841 TGTTCATGTTTAAACGGTAAGTCAATTCGCCAGCTAGAGACGACCATCAAGTAA 900
Db 852 TGTTCATGTTTAAACGGTAAGTCAATTCGCCAGCTAGAGACGACCATCAAGTAA 911
OY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGCTGGGGA 949
Db 912 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGCTGGGGA 960

RESULT 4
US-08-600-993A-3
; Sequence 3, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; NUMBER OF SEQUENCES: 59
; TITLE OF INVENTION: CELLULAR VACCINES BASED THEREON
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 960 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:
; CLONE: SPRU42
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..960
; US-08-600-993A-3

Query Match 47.2%; Score 944.2; DB 2: Length 960;
Best Local Similarity 99.7%; Pred. No. 1.9e-278;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TAAATCTACGACAAATTAATCACTCATTTGCTGATGGGTTGAAAGCCGCTCA 60
Db 12 TAAATCTACGACAAATTAATCACTCATTTGCTGATGGGTTGAAAGCCGCTCA 71
OY 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATGCTTATGAGACCA 120
Db 72 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATGCTTATGAGACCA 131
OY 121 TCGCTTCTCGACACAGAGGGGATTTGATACACCTCTATCCTGGAGCTTTCTGGGCA 180
Db 132 TCGCTTCTCGACACAGAGGGGATTTGATACACCTCTATCCTGGAGCTTTCTGGGCA 191
OY 181 TCTGCAAGCAATTCCTCCAAAGTGATCACTCCACCAACAGTGAATTAAGTTGAC 240
Db 192 TCTGCAAGCAATTCCTCCAAAGTGATCACTCCACCAACAGTGAATTAAGTTGAC 251
OY 241 TTACTTTTCACTTCGACTTCGACCACTATTTCTGTTAAGGCTCAGAGAGCTTGGTT 300
Db 252 TTACTTTTCACTTCGACTTCGACCACTATTTCTGTTAAGGCTCAGAGAGCTTGGTT 311
OY 301 AGCGATTCACTTAGAACAACCAAGCAAGCAAGAAATCTGACTACTATTAATTA 360
Db 312 AGCGATTCACTTAGAACAACCAAGCAAGCAAGAAATCTGACTACTATTAATTA 371
OY 361 GGTCTACATGCTCAATGGGAACATGATGACAGACGACGCTCAAACTACTATGTTAA 420
Db 372 GGTCTACATGCTCAATGGGAACATGATGACAGACGACGCTCAAACTACTATGTTAA 431
OY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCTTGTGGTGAATGCTTCAGGCACC 480
Db 432 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCTTGTGGTGAATGCTTCAGGCACC 491
OY 481 AAACCAATATGACCCCTATTCACATCCAGAGACCCCAAGACCCGCCGAACTTGGTCT 540
Db 492 AAACCAATATGACCCCTATTCACATCCAGAGACCCCAAGACCCGCCGAACTTGGTCT 551
OY 541 ATCTGAATGAACAAATCAAGGGCTACATCTCTGTGAACGATATGAAGAACACTCAATAC 600
Db 552 ATCTGAATGAACAAATCAAGGGCTACATCTCTGTGAACGATATGAAGAACACTCAATAC 611
OY 601 ACCAATTACTGATGACATCAAAAGTCTCAATTCAGCAAGTAATTAACCTGCTTACATGA 660
Db 612 ACCAATTACTGATGATGACATCAAAAGTCTCAATTCAGCAAGTAATTAACCTGCTTACATGA 671
OY 661 TAATTACCTCAAGGAAGTATCATCACTTAAGAGAAAGAGCTATTAACCTACTGC 720
Db 672 TAATTACCTCAAGGAAGTATCATCACTTAAGAGAAAGAGCTATTAACCTACTGC 731
OY 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
Db 732 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 791
OY 781 TTACATACAGACGAATACGTTGCCATACAGACGATGAATTCGAATGCTTCTACCAT 840
Db 792 TTACATACAGACGAATACGTTGCCATACAGACGATGAATTCGAATGCTTCTACCAT 851

Db 200 GCTTCTACGACATCAGCGGTTGACCCGGTGGGATCTTCCTGACGACAGCCGTGGCGC 259
QY 183 TGCAGCAATTCCTCCAGGATGATGATCACTCTCACCCAACTGATTAAGTTGACTT 242
Db 260 TGTTCCTCCGTCACGCGTCACAGGGCACTACCATTTACCCACAGCTGGCGGAAACT 319
QY 243 ACTTTTCACTTCGACTTCGACGACGACTATTTCTCGTAAGGCTCAGGAACCTTGTTAG 302
Db 320 TCTTCCTCAGTCC-----AGAACGCAAGCTGATGCGTAAGATTAAGGAAGCTTCTCCG 373
QY 303 CGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACTACTATTAATAAGG 362
Db 374 CGATTTCAGTTAGAACAACTGCTGACGAAAGAGATCTCGACTTATCTGAAAGAGA 433
QY 363 TCTCATGCTTAATGGAAGATGATGATGACAGACGAGCTCAAACTACTATGTAAG 422
Db 434 TTTACCTTGTTACCGGCGCTATGCTGCTGCGGCAAGCTATTTCCGAAAAA 493
QY 423 ACCTCAATTAATTTAGTTTACCTGAGTTAGCCTTGCTGCTGGAATGCTCAGGACCAA 482
Db 494 CGGTGACCAACTGACGCTGAGCAAAATGGGGTATAGCGGGCTGCGCAAGCGCCTT 553
QY 483 ACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGCCGGAATGCTGTTAT 542
Db 554 CCACCTTCAACCCCTCTACTGATGATGCTGCGCTGCGGCTGTAACCTGCTGCTGT 613
QY 543 CTGAATAGAAAATCAAGGCTACATCTGCTGAACAGATGAGAAAGCAAGTCAATAAC 602
Db 614 CGCGGATGCTGAGTAAGGATATATACCCCAACAGTTGATCAGACGACGACTGAGG 673
QY 603 CAATTACTGAT 613
Db 674 CGATTACGCT 684

RESULT 8
US-08-771-716-1
; Sequence 1, Application US/08771716
; Patent No. 5922540
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, Joann
; APPLICANT: Jaskunas, S.Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase
; TITLE OF INVENTION: Gene of Staphylococcus Aureus
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,716
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEO ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..807
US-08-771-716-1

Query Match 3.18; Score 62.2; DB 2; Length 807;
Best Local Similarly 49.9%; Pred. No. 3.4e-09;
Matches 191; Conservative 0; Mismatches 183; Indels 9; Gaps 1;

QY 67 AGCTAATGATATATCCACAGATGTTGGTTAAGGCAATGCTTCTATCGAAGCATTCCTT 126
Db 252 AGCTGATTAACATGCGAGAGATGTTAAAGTGCTTTATTTCAATGGAAGTGAACGATT 311
QY 127 CTTCGACCACAGGGGATGATGATACATCCGTATCCCTGGAGCTTCTTGGCAATTCGCA 186
Db 312 CTCAATATCATATGATGATGATTTGAAAGTACAACTAGACCTTATTTCAACGATTAG 371
QY 187 AAGCAATTCCTCCAGGTGATCACTTCACCCAAAGTTGATTAAGTGAATCTTACT 246
Db 372 CGACAGAGATGTCAGAGGTGATGATACATTAACAAACAAAGTTGCAAAATTAATTTT 431
QY 247 TTCACTTCGACTTCGACAGCATATTTCTGTAAGGCTGAGAGAGCTGTTAGGAT 306
Db 432 TGATTAAT-----GATGCTTATTACTGAAAGAAATTAATTTTGTAGTCA 482
QY 307 TCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATTAATAAGGCTTA 366
Db 483 TCGAGTTGAAAACAAATATATATAGAACGAATTTTAAAGTTTATTAATAATATTTA 542
QY 367 CATGCTAATGGAAGTATGATGATGACAGACGAGCTCAAACTACTATGTAAGACT 426
Db 543 CTTGGGATTAATCAATATATAGCTTGAAGGGCGAGCAAAACATTACTTGGAAACACGT 602
QY 427 CAATAATTTAGTTTACCTGAT 449
Db 603 GAATTAATAATGTAACACATGT 625

RESULT 9
US-08-771-716-3
; Sequence 3, Application US/08771716
; Patent No. 5922540
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, Joann
; APPLICANT: Jaskunas, S.Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase
; TITLE OF INVENTION: Gene of Staphylococcus Aureus
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,716
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:


```
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,720A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11067
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 807 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-057-720A-3
```

```
Query Match 3.1%; Score 62.2; DB 3; Length 807;
Best Local Similarity 33.9%; Pred. No. 3.4e-09;
Matches 130; Conservative 61; Mismatches 183; Indels 9; Gaps 1;

QY 67 AGCTAATGATTTCCACAGATTGGTTAAGCAATCGTTTCTTCGACGACCATCGCTT 126
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 AGCGAUAACAUGCCAGAGUAGUUAAGGUGCCUUUUAUUGCAUUGAAGAUACAAGAU 311
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 CTTCGACCACGAGGGGATTTACCATCCGTATCCTGGAGCTTCTGGCCATCTGCA 186
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 312 CUACAUAUCAUGAGAUUCGUAUUUAGAAAGUACAUAAGCUUUUAUUCACAGAUAG 371
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 AAGCAATTCCTCCCAAGGTGATCACTTCACCCACAGTTGATTAAGTTAGTTACTT 246
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 372 CGACAGAGAUUGCAGUGUGUAGUACAUUACAACAAGUUGUCAAUAUUUUUA 431
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 247 TTCACCTTCGACTTCGACACGACTATTTCGTGAGGCTCAGAGCTTGGTTAGCAT 306
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 432 UGAUAU-----GATCCUDCAUUDACUAGAAAGUAAAGAUUUAUUGUAGCUCA 482
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 307 TCAGTTGAACAAAAGCAACCAAGAAATCTTGACCTATATATAATAAGGTCTA 366
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 483 UCGAGUUGAUAUAUAUAUAAGAGAAUUAAGCUUUUAUUAUUAUUAUUA 542
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 367 CATCTTAATGGGAATGAGATGACAGACAGCTCAAAACTACTATGTTAAAGACT 426
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 543 CUUUGGGAUUAUCAUUAUAGCUUGAGGGCGACGAAACAUUAUUUGAACAACCGU 602
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 427 CAATAATTTAAGTTTACCTCAGT 449
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 603 GAATAAAUUAGUACAACAAUGU 625
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 12
US-08-731-716-1
Sequence 1, Application US/08731716
Patent No. 5789202
GENERAL INFORMATION:
APPLICANT: Hoskins, Joan
APPLICANT: Jaskunas, S. Richard
APPLICANT: Rocky, Pamela K.
```

```
APPLICANT: Zhao, Genshi
APPLICANT: Rosteck, Paul R. Jr.
APPLICANT: No. 5789202ris, Franklin H.
TITLE OF INVENTION: Penicillin Binding Protein From
TITLE OF INVENTION: Streptococcus Pneumoniae
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731,716
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-10,887
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2193 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2193
US-08-731-716-1
```

```
Query Match 2.8%; Score 56.4; DB 1; Length 2193;
Best Local Similarity 44.6%; Pred. No. 3.4e-07;
Matches 381; Conservative 0; Mismatches 446; Indels 27; Gaps 3;

QY 5 ATCTACGACATATAAATCACTCATTTGGGTGTAAGCCGCGCTCAATGCC 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 283 ATTTTGAACCGTGAAGAAAGAGCGTGTCCTGTCTGTCAAAAGGAACCTATGTT 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 CAAGCTATATGATTTCCACAGATTTGGTTAAGCAATCGTTTATGAGAACCATCGC 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 343 GAGCTGACTGCATCAAGTAAACTTGCAGAAATGCTGTATTTGGACAGAAAGCGTTCT 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 TTCTTCGACCAAGGGGATTTGATACATCCGTAATCTGGAGGCTTCTTGGCAATCTG 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 403 TTCTATAAATAATGACGGGATTAACATAGGCCGTTTC-----TTCTTGGCTATTGTC 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 CAAGCAATTCCTCCCAAGTGAATCAACTCTCACCCACAGTTGATTAAGTTACTTAC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 454 ACTGCTGAGCGTTGAGGTGGCTCTACCATTTACCAACAGCTGCTAAAMACGCCAT 513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 245 TTTTCAACTTGACTTCGACCAAGATTTCTGTAAGGCTCAGAAAGCTTGTTAGCG 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 514 TTAT-----CGAGGATCAAACTGTGTAGAGAAAGCGAAAGATTTTCTCTTGCC 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 305 ATTCAAGTTAGCAAAAAGCAACCAAGCAAGAAATCTTGACTACTATTAATAAGTGC 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 565 TTAGATTTAAGCAAAAATAATAGTAAGAGCAATTTCTAACCATCTTAACCAAGCT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 TACATGTCTAATGGGAACCTATGAGATGACAGACAGCTCAAAACTACTATGTTAAAGAC 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 12:37:03 ; Search time 3441.19 Seconds
(without alignments)
12567.793 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTAGACATATAA.....ATCCTACACGACACAACA 1999

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

- 1: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
- 2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
- 3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
- 4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
- 5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
- 6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
- 7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
- 8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
- 9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
- 10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
- 11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
- 12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
- 13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
- 14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
- 15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
- 16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
- 17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
- 18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
- 19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
- 20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
- 21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
- 22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
- 23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
- 24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
- 25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
- 26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
- 27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
- 28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
- 29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
- 30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
- 31: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
- 32: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
- 33: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
- 34: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
- 35: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
- 36: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
- 37: /cgn2_6/ptodata/2/pna/US100_COMB.seq:*
- 38: /cgn2_6/ptodata/2/pna/US101_COMB.seq:*
- 39: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
- 40: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
- 41: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
- 42: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
- 43: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*

- 44: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
- 45: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
- 46: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
- 47: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
- 48: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
- 49: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
- 50: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
- 51: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
- 52: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
- 53: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
- 54: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
- 55: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
- 56: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
- 57: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
- 58: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
- 59: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
- 60: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
- 61: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
- 62: /cgn2_6/ptodata/2/pna/US6023_COMB.seq:*
- 63: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
- 64: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
- 65: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
- 66: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
- 67: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
- 68: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
- 69: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
- 70: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
- 71: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
- 72: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
- 73: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
- 74: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
- 75: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	20 US-09-536-784-1	Sequence 1, Appli
2	1999	100.0	1999	30 US-09-765-271-1	Sequence 1, Appli
3	1999	100.0	1999	30 US-09-765-272-1	Sequence 1, Appli
4	1999	100.0	10383	41 US-60-029-960-81	Sequence 145, App
5	1999	100.0	10711	13 US-08-961-527-145	Sequence 112, Ap
6	1987.8	99.4	2160	22 US-09-583-110-1312	Sequence 9325, Ap
7	1987.8	99.4	2166	15 US-09-107-433-1102	Sequence 1102, Ap
8	1981.4	99.1	2160	31 PCR-US02-03987-9325	Sequence 9325, Ap
9	1981.4	99.1	2160	31 US-09-815-242-9325	Sequence 9325, Ap
10	1981.4	99.1	2160	37 US-10-072-851-9325	Sequence 596, App
11	1981.4	99.1	10333	45 US-60-061-998-596	Sequence 596, App
12	1631.2	81.6	9845	45 US-60-068-175-596	Sequence 139, Appli
13	944.2	47.2	960	5 US-08-116-541-3	Sequence 139, Appli
14	808.6	40.5	2172	13 PCR-US97-14436-139	Sequence 139, App
15	808.6	40.5	2172	13 US-08-911-503-139	Sequence 139, App
16	808.6	40.5	2172	13 US-08-911-503A-139	Sequence 139, App
17	574.6	28.7	2337	1 PCR-US02-03987-6631	Sequence 6631, App
18	574.6	28.7	2337	31 US-09-815-242-6631	Sequence 6631, App
19	574.6	28.7	2337	37 US-10-072-851-6631	Sequence 6631, App
20	573	28.7	2370	15 US-09-134-000-1534	Sequence 1534, App
21	559	28.0	8395	14 US-09-070-927-217	Sequence 217, App
22	559	28.0	8395	14 US-09-070-927A-217	Sequence 217, App
23	553	27.7	2472	15 US-09-107-532-2013	Sequence 2013, App
24	553	27.7	2472	15 US-09-107-532A-2013	Sequence 2013, App
25	496.2	24.8	1386	37 US-10-091-007-91	Sequence 495, Appli
26	410	20.5	2470	44 US-60-068-186-495	Sequence 495, App
27	410	20.5	2474	44 US-60-050-444-495	Sequence 495, App
28	352.2	17.6	3478	45 US-60-068-217-710	Sequence 710, App
29	348.6	17.4	3279	44 US-09-634-228-44	Sequence 44, Appli
30	348.6	17.4	7818	24 US-09-634-238-14	Sequence 14, Appli
31	273	13.7	933	43 US-60-045-649-108	Sequence 108, App

1201 AATGCACTACCCAGATATCTACTACTCAATGCAATTCGAACTAACACCAACGAATCAGA 1260
1201 AATGCACTACCCAGATATCTACTACTCAATGCAATTCGAACTAACACCAACGAATCAGA 1260
1261 CAAAAAATATGAGCAAGTAGTGAAGAAAGATGCTGCTGCTTACGCTTGGCAATGG 1320
1261 CAAAAAATATGAGCAAGTAGTGAAGAAAGATGCTGCTGCTTACGCTTGGCAATGG 1320
1321 TGGAACTTACTTAATACCAATATATATCATATTAAGTCGCTTATGATGAGGAGTGAAG 1380
1321 TGGAACTTACTTAATACCAATATATATCATATTAAGTCGCTTATGATGAGGAGTGAAG 1380
1321 TGGAACTTACTTAATACCAATATATATCATATTAAGTCGCTTATGATGAGGAGTGAAG 1380
1381 AGAGTTCTCTAATGTCGCAACTGCGCCATGATGAAGAAAGACAGCCCTATATGATGACCA 1440
1381 AGAGTTCTCTAATGTCGCAACTGCGCCATGATGAAGAAAGACAGCCCTATATGATGACCA 1440
1381 AGAGTTCTCTAATGTCGCAACTGCGCCATGATGAAGAAAGACAGCCCTATATGATGACCA 1440
1441 CATGATGAAAACAGTCTTGTACTTATGAACTGAGGAAATGCTTATGCTTGGCTGCC 1500
1441 CATGATGAAAACAGTCTTGTACTTATGAACTGAGGAAATGCTTATGCTTGGCTGCC 1500
1501 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1560
1501 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1560
1561 GACCTCTCAATTTGTAGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1561 GACCTCTCAATTTGTAGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1621 GGCCTGATGAGAGGCTATTTTACCGTCTGACACCACTTGTAGGCAATGGCTTACGCT 1680
1621 GGCCTGATGAGAGGCTATTTTACCGTCTGACACCACTTGTAGGCAATGGCTTACGCT 1680
1681 CGCTGCCAAGTTTACCGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1681 CGCTGCCAAGTTTACCGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1741 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1800
1741 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1800
1741 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1800
1801 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1860
1801 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1860
1861 ATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1861 ATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1921 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1980
1921 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1980
1981 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1980
1981 TTAGGCTGGTAAACAGGAACTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1980

RESULT 2
US-09-765-271-1
Sequence 1, Application US/09765271
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage

COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,271
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/536,784
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1999 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-765-271-1

Query Match 100.0%; Score 1999; DB 30; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TAAATCTACGACAAATTAATCACTATGCTGATGCTGGTCTGAAAGCCGCTCAA 60
1 TAAATCTACGACAAATTAATCACTATGCTGATGCTGGTCTGAAAGCCGCTCAA 60
61 TGGCCAGGTAATGATATTCACAGATTTGTTAAGGCAATCGTTTATGCAAGCA 120
61 TGGCCAGGTAATGATATTCACAGATTTGTTAAGGCAATCGTTTATGCAAGCA 120
121 TGGCTCTTCGACACAGGAGGATGATACATCCGATCTGAGAGCTTCTTGGGCAA 180
121 TGGCTCTTCGACACAGGAGGATGATACATCCGATCTGAGAGCTTCTTGGGCAA 180
181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTACCCCAACAGTGAATTAAGTTGAC 240
181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTACCCCAACAGTGAATTAAGTTGAC 240
241 TTAGCTTTTCACTGACCTTCGACCACTATTTCTGTAAGGCTCAGAGAGCTTGGTT 300
241 TTAGCTTTTCACTGACCTTCGACCACTATTTCTGTAAGGCTCAGAGAGCTTGGTT 300
241 TTAGCTTTTCACTGACCTTCGACCACTATTTCTGTAAGGCTCAGAGAGCTTGGTT 300
301 AGGATTCAGTTAGAACCAAAAGCAACAGCAAGCAAAATCTGACCTACTATTAATAA 360
301 AGGATTCAGTTAGAACCAAAAGCAACAGCAAGCAAAATCTGACCTACTATTAATAA 360
301 AGGATTCAGTTAGAACCAAAAGCAACAGCAAGCAAAATCTGACCTACTATTAATAA 360
361 GGTCTACATGCTTAATGGAAGTATGATGATGATGATGATGATGATGATGATGATGAT 420
361 GGTCTACATGCTTAATGGAAGTATGATGATGATGATGATGATGATGATGATGATGAT 420
421 AGACCTCAATTAATTAAGTTTACCTGATGATGATGATGATGATGATGATGATGATGAT 480
421 AGACCTCAATTAATTAAGTTTACCTGATGATGATGATGATGATGATGATGATGATGAT 480
481 AAACCAATATGACCCCTATTCACATGCAAGAGCCCAAGAGCCGCGAACTTGGTCT 540
481 AAACCAATATGACCCCTATTCACATGCAAGAGCCCAAGAGCCGCGAACTTGGTCT 540
541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAAGAGTATGAAAGAGAGTCAATAC 600
541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAAGAGTATGAAAGAGAGTCAATAC 600

QY	601	ACCAATTAACGATGAGACTATCAAAAGTCTCAAAATCAGCAAACTAATTAACCTGGCTACATATGA	660
Db	601	ACCAATTAACGATGAGACTACAAAAGTCTCAAAATCAGCAAACTAATTAACCTGGCTACATATGA	660
QY	661	TAATTAACCTCAAGGAAGTCTCATCAATCAAGTTGGAAGAAGCAAGGCTATTAACCTACTCAC	720
Db	661	TAATTAACCTCAAGGAAGTCTCATCAATCAAGTTGGAAGAAGCAAGGCTATTAACCTACTCAC	720
QY	721	AACCTGGATGGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
Db	721	AACGGGATGGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
QY	781	TTTCAATACAGACGAATAAGCTGTGCATTCAGACGATGAAATTTGCAAGTGGCTCTACCAT	840
Db	781	TTTCAATACAGACGAATAAGCTGTGCCTTATCCAGACGATGAAATTTGCAAGTGGCTCTACCAT	840
QY	841	TGTTGATGTTTCTTAACGCTAAAGTCTATTCGCCACGCTAGGAGCAGCCATCAGTCAAGTAA	900
Db	841	TGTTGATGTTTCTTAACGCTAAAGTCTATTCGGCCAGCTAGGAGCAGCCATCAGTCAAGTAA	900
QY	901	TGTTTCTTGGGAATTACCAAGCACTAGAACAAACCCGACTGGGATCACTATGAA	960
Db	901	TGTTTCTTGGGAATTACCAAGCACTAGAACAAACCCGACTGGGATCACTATGAA	960
QY	961	ACCGATACAGACTATCTCTGCTGGAGTACGGTGTCTAGATTAACAGTCACTACT	1020
Db	961	ACCGATACAGACTATCTCTGCTGGAGTACGGTGTCTAGATTAACAGTCACTACT	1020
QY	1021	CGTTACAGATAGAGCCCTTAATCACTACCCCTGGAGCAAAATACCTCTGTTTAACTGGGATAG	1080
Db	1021	CGTTACAGATAGAGCCCTTAATCACTACCCCTGGAGCAAAATACCTCTGTTTAACTGGGATAG	1080
QY	1081	GGGCTACTTTGGCAATCATACCTTGCAATTAACGCTTCGAACATGCGGAAAGTCCAGC	1140
Db	1081	GGGCTACTTTGGCAATCATACCTTTGCAAAATACGCTTCGAACATGCGGAAAGTCCAGC	1140
QY	1141	CGTGGAAACTCTAAACAAGGTGGGACTCAACCGCCGACAGACTTTCTCAATAGTGTACG	1200
Db	1141	CGTGGAAACTCTAAACAAGGTGGGACTCAACCGCCGACAGACTTTCTCAATAGTGTACG	1200
QY	1201	AATGACTACCCAAAGTATTAATCACTACTCAATGCGATTTCAATTAACACACCGAATCGA	1260
Db	1201	AATGACTACCCAAAGTATTAATCACTACTCAATGCGATTTCAATTAACACACCGAATCGA	1260
QY	1261	CAAAAAATATGAGACAGTAGTAGAAAAAGATGGCTGTGCTACGCTGGCTTGGCAATGG	1320
Db	1261	CAAAAAATATGAGACAGTAGTAGAAAAAGATGGCTGTGCTACGCTGGCTTGGCAATGG	1320
QY	1321	TGGAACTACTATTAACCAATGTATATGCCATTAAGTCGCTTTACTGATGGAGAGTGA	1380
Db	1321	TGGAACTACTATTAACCAATGTATATGCCATTAAGTCGCTTTACTGATGGAGAGTGA	1380
QY	1381	AGAGTTCTATATGTGCGAAGCTGTCACATGGAAGAAAGACAGCGCTATATGATAGCGGA	1440
Db	1381	AGAGTTCTCTAATGTGCGAAGCTGTGTCATGGAAGAAAGACAGCGCTATATGATAGCGGA	1440
QY	1441	CATGATGAAAAACAGTCTTGACTTATGGAACATGGACAAATGCCATCTTGCTTGGCTGCC	1500
Db	1441	CATGATGAAAAACAGTCTTGACTTATGGAACATGGACAAATGCCATCTTGCTTGGCTGCC	1500
QY	1501	TCGAGCTGTGTAAACAGGAACCTTAACATATACAGACGAGGAATTTGAAACCAATCAA	1560
Db	1501	TCGAGCTGTGTAAACAGGAACCTTAACATATACAGACGAGGAATTTGAAACCAATCAA	1560
QY	1561	GACCTCAATTTGTATACACCTGATGACATATTTGGTGGCTTTACGCGTAATATTCAT	1620
Db	1561	GACCTCTCAATTTGTATACACCTGATGACATATTTGGTGGCTTTACGCGTAATATTCAT	1620
QY	1621	GGCTGTATGACAGAGCTATTTCAACGCTGACACCACTTTAGGCAATGGCTTACGCT	1680
Db	1621	GGCTGTATGACAGAGCTATTTCAACGCTGACACCACTTTAGGCAATGGCTTACGCT	1680
QY	1681	CGCTGCCAAAGTTTACCGCTCTGTATGACTTAACCTGTCTGAAGAGACAAATCCAGAAGA	1740

Db	1681	CGCGCCAAAGTTACCGGCTCTATGATGACCTACGTGTCTGAAGGAGCAATCCAGAA	17400
Qy	1741	TTGGAATATATACAGAGGGGCTTACAGAAATGAGAAATTCGATTTTAAAAATGGTGCTCG	18000
Db	1741	TTGGAATATATACAGAGGGGCTTACAGAAATGAGAAATTCGATTTTAAAAATGGTGCTCG	18000
Qy	1801	TTTACGCGGAAACACCTGCTCCACAAACCCCATCACTACATGAAAGTTCAAAGCTCATC	18666
Db	1801	TTTACGCGGAAACACCTGCTCCACAAACCCCATCACTACATGAAAGTTCAAAGCTCATC	18666
Qy	1861	ATCAGATAGTTCAACTTCACAGCTCTAGCTCAACCACTCCAGACACAATAATATAGTAGAC	19200
Db	1861	ATCAGATAGTTCAACTTCACAGCTCTAGCTCAACCACTCCAGACACAATAATATAGTAGAC	19200
Qy	1921	TACCAATCTTACCAATATATACGCAACCAATTAATACAAACCCTGATCAACAAATTCAGAA	19800
Db	1921	TACCAATCTTACCAATATATACGCAACCAATTAATACAAACCCTGATCAACAAATTCAGAA	19800
Qy	1981	TCCTCAACGACACACACCA 1999	
Db	1981	TCCTCAACGACACACACCA 1999	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Db 1 TAAAACTACGACAAATAAAAATCAATCATTTGCTGACTTGGTGTCTGAAGCCGCGCTCAA 60
 QY 120 TCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGTTTATGGAAGACCA 120
 Db 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGTTTATGGAAGACCA 120
 QY 121 TCGCTTCTCGACACAGAGGGGGATGATACATCCGATCTCGGAGCTTTCTGCGCAA 180
 Db 121 TCGCTTCTCGACACAGAGGGGGATGATACATCCGATCTCGGAGCTTTCTGCGCAA 180
 QY 181 TCTGCAAAAGCAATTCCTTCACAGGTGGATCAACTCTCACCAACAGTTGATTAAGTTGAC 240
 Db 181 TCTGCAAAAGCAATTCCTTCACAGGTGGATCAACTCTCACCAACAGTTGATTAAGTTGAC 240
 QY 241 TTACCTTCACTTCGACCTCCGACAGACTTTTCTCTGTAAGGCTCAGGAAGCTTGGT 300
 Db 241 TTACCTTCACTTCGACCTCCGACAGACTTTTCTCTGTAAGGCTCAGGAAGCTTGGT 300
 QY 301 AGCGATTGAGTTAGCAAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATATTA 360
 Db 301 AGCGATTGAGTTAGCAAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATATTA 360
 QY 361 GGTCTACATGTCTTAATGGGAATGGAATGCAAGCAGCTCAAAACTACTATGTGTA 420
 Db 361 GGTCTACATGTCTTAATGGGAATGGAATGCAAGCAGCTCAAAACTACTATGTGTA 420
 QY 421 AGACCTCAATAATTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGCTCAGGACAC 480
 Db 421 AGACCTCAATAATTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGCTCAGGACAC 480
 QY 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGCCGCAAACTTGGTCT 540
 Db 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGCCGCAAACTTGGTCT 540
 QY 541 ATCTGAAAATGAAAAATCAAGGCTACATCTCTGCTGAAGAGTGAAGAAAGCAGTCAATAC 600
 Db 541 ATCTGAAAATGAAAAATCAAGGCTACATCTCTGCTGAAGAGTGAAGAAAGCAGTCAATAC 600
 QY 601 ACCATTCTGATGAGTCAACAAAGTCAATCAGCAAGTATTAACCTGCTTAACATGA 660
 Db 601 ACCATTCTGATGAGTCAACAAAGTCAATCAGCAAGTATTAACCTGCTTAACATGA 660
 QY 661 TAATTACCTCAAGAAAGTCAATCAAGTTGAGAAAGAAAGAGGCTTAACCTACTCAC 720
 Db 661 TAATTACCTCAAGAAAGTCAATCAAGTTGAGAAAGAAAGAGGCTTAACCTACTCAC 720
 QY 721 AACTGGGATGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
 Db 721 AACTGGGATGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
 QY 781 TTACATATCAGACGAATACGTTGCTATCAGAGCATGATGCAAGTCTGTTACCAT 840
 Db 781 TTACATATCAGACGAATACGTTGCTATCAGAGCATGATGCAAGTCTGTTACCAT 840
 QY 841 TGTTCATGTCTTAACGGTAAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Db 841 TGTTCATGTCTTAACGGTAAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 QY 901 TGTTCATGTCTTAACGGTAAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 Db 901 TGTTCATGTCTTAACGGTAAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 QY 961 ACCGATCAGACATATGCTTCCGCTTGGAGTACGCTGTACATGTTCAACTGCTACTAT 1020
 Db 961 ACCGATCAGACATATGCTTCCGCTTGGAGTACGCTGTACATGTTCAACTGCTACTAT 1020
 QY 1021 CGTTACAGATGAGCCCTTAATACCTGAGCAATATCTCTGTTTAACTGGGATATG 1080
 Db 1021 CGTTACAGATGAGCCCTTAATACCTGAGCAATATCTCTGTTTAACTGGGATATG 1080
 QY 1081 GGGCTACTTTGGCAACATACCTTGCATATAGCCCTGCAACAATGCGCAAAAGCTCCAGC 1140
 Db 1081 GGGCTACTTTGGCAACATACCTTGCATATAGCCCTGCAACAATGCGCAAAAGCTCCAGC 1140

QY 1141 CGTGGAACTTAACAAGTGCAGACTCAACCGCGCAAGACTTTCTTAATGTGTAGG 1200
 Db 1141 CGTGGAACTTAACAAGTGCAGACTCAACCGCGCAAGACTTTCTTAATGTGTAGG 1200
 QY 1201 AATCGACTACCCAGATTTCACTACTCAAAATGCCATTTCAAGTAAACACACGAATCAGA 1260
 Db 1201 AATCGACTACCCAGATTTCACTACTCAAAATGCCATTTCAAGTAAACACACGAATCAGA 1260
 QY 1261 CAAAAAATATGAGCAAGTATGAAAAAGATGCTGCTTACGCTGCTTTGCAATATG 1320
 Db 1261 CAAAAAATATGAGCAAGTATGAAAAAGATGCTGCTTACGCTGCTTTGCAATATG 1320
 QY 1321 TGGACCTTCTTAACCAATGATATATCATTAAGACGCTTTAGAGAGGAGTGA 1380
 Db 1321 TGGACCTTCTTAACCAATGATATATCATTAAGACGCTTTAGAGAGGAGTGA 1380
 QY 1381 AGAGTTCTCTAATGTGCAAGTGCAGCTGATGCAAGGAAAGCAAGCCTATATGATGACGA 1440
 Db 1381 AGAGTTCTCTAATGTGCAAGTGCAGCTGATGCAAGGAAAGCAAGCCTATATGATGACGA 1440
 QY 1441 CATGATGAAAACAGTCTTGACTTATGGAACCTGAGCAAAATGCTTGTGCTGCC 1500
 Db 1441 CATGATGAAAACAGTCTTGACTTATGGAACCTGAGCAAAATGCTTGTGCTGCC 1500
 QY 1501 TCGAGCTGTAAACAGAGACCTCTACTATACAGACGAGGAAATTTGAACACATCA 1560
 Db 1501 TCGAGCTGTAAACAGAGACCTCTACTATATACAGACGAGGAAATTTGAACACATCA 1560
 QY 1561 GACCTCTCAATTTGAGACCTGATGAACTATTTGCTGCTTATACGCGTAATATTCAT 1620
 Db 1561 GACCTCTCAATTTGAGACCTGATGAACTATTTGCTGCTTATACGCGTAATATTCAT 1620
 QY 1621 GCGTGTATGAGAGGCTATTTACCGTCTGACACCACTTGTGAAGGAAGCAATCCAGAA 1740
 Db 1621 GCGTGTATGAGAGGCTATTTACCGTCTGACACCACTTGTGAAGGAAGCAATCCAGAA 1740
 QY 1681 GCGTGTATGAGAGGCTATTTACCGTCTGATGACCTGCTGTAAGGAAGCAATCCAGAA 1800
 Db 1681 GCGTGTATGAGAGGCTATTTACCGTCTGATGACCTGCTGTAAGGAAGCAATCCAGAA 1800
 QY 1741 TTGGAATATACAGAGGAGGCTCTACAGAAATGAGAAATGCTTAATTAATGAGTCTG 1800
 Db 1741 TTGGAATATACAGAGGAGGCTCTACAGAAATGAGAAATGCTTAATTAATGAGTCTG 1800
 QY 1801 TTTCTAGTGAAGTCACTGCTCAACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
 Db 1801 TTTCTAGTGAAGTCACTGCTCAACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
 QY 1861 ATCAGATAGTTCAACTTCAAGTCTGACAGTCAACCACTCCAGCACCAATATATGATGAC 1920
 Db 1861 ATCAGATAGTTCAACTTCAAGTCTGACAGTCAACCACTCCAGCACCAATATATGATGAC 1920
 QY 1921 TACCAATCTTAACAATATATGCAACCAATCAATCAATCAATCAATCAATCAATCA 1980
 Db 1921 TACCAATCTTAACAATATATGCAACCAATCAATCAATCAATCAATCAATCAATCA 1980
 QY 1981 TCCTCAACGACCAACCA 1999
 Db 1981 TCCTCAACGACCAACCA 1999

RESULT 4
 US-60-029-960-81/c
 : Sequence 81, Application US/60029960
 : GENERAL INFORMATION:
 : APPLICANT: Charles Kunsch
 : TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 : NUMBER OF SEQUENCES: 1649
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Human Genome Sciences, Inc.
 : STREET: 9410 Key West Avenue
 : CITY: Rockville

OY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGGAAGAGCAATCCAGAA 1740
|||||
Db 8083 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGGAAGAGCAATCCAGAA 8024
OY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGCTGCG 1800
|||||
Db 8023 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGCTGCG 7964
OY 1801 TTCTACGTTGAAGTACCTGCTGTCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 1860
|||||
Db 7963 TTCTACGTTGAAGTACCTGCTGTCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 7904
OY 1861 ATCAATGTTGTAATCTACAGTCTAGCTCAACCACTCCAAACCAAAATATGATGAC 1920
|||||
Db 7903 ATCAATGTTGTAATCTACAGTCTAGCTCAACCACTCCAAACCAAAATATGATGAC 7844
OY 1921 TACCAATCTTAACAATATACGCAACAATCAATCAACCCCTGATCAACAATACAGA 1980
|||||
Db 7843 TACCAATCTTAACAATATACGCAACAATCAATCAACCCCTGATCAACAATACAGA 7784
OY 1981 TCCTCAACGACGACCA 1999
|||||
Db 7783 TCCTCAACGACGACCA 7765

RESULT 5
US-08-961-527-145/c
; Sequence 145, Application US/08961527
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8512
; TELEFAX: (301) 309-8504
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-145

Query Match 100.0%; Score 1999; DB 13; Length 10711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAAATCTACGACATATAAATCACTCATTTGCTGACTTGGGTTCTGAACGGCGGTCA 60

Db 9767 TAAATCTACGACATATAAATCACTCATTTGCTGACTTGGGTTCTGAACGGCGGTCA 9708
OY 61 TCCCAAGCTAATGATATATCCACAGATTTGGTTAAGCAATCGTTTCTATGAGACCA 120
|||||
Db 9707 TCCCAAGCTAATGATATATCCACAGATTTGGTTAAGCAATCGTTTCTATGAGACCA 9648
OY 121 TGGCTTCTGACACAGGGGGATGATACATCCGTATCCGGGAGCTTCTTCCGCA 180
|||||
Db 9647 TGGCTTCTGACACAGGGGGATGATACATCCGTATCCGGGAGCTTCTTCCGCA 9588
OY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCCACCCAAAGTTGATTAAGTTGAC 240
|||||
Db 9587 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCCACCCAAAGTTGATTAAGTTGAC 9528
OY 241 TTACTTTTCACTTGACCTGCTCGACAGACTATTTCTGTAAAGCTCAGAAAGCTTGGTT 300
|||||
Db 9527 TTACTTTTCACTTGACCTGCTCGACAGACTATTTCTGTAAAGCTCAGAAAGCTTGGTT 9468
OY 301 AGCGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTGACCTACTATATAATA 360
|||||
Db 9467 AGCGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTGACCTACTATATAATA 9408
OY 361 GGTCTACATGTCTAATGGAAGTATGGAATGAGACAGCAGCTCAAACTACTATAGTAA 420
|||||
Db 9407 GGTCTACATGTCTAATGGAAGTATGGAATGAGACAGCAGCTCAAACTACTATAGTAA 9348
OY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGTGGCTGGAATCCTCAGGACCC 480
|||||
Db 9347 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGTGGCTGGAATCCTCAGGACCC 9288
OY 481 AAACCAATATGACCCCTATTCATTCACCAAGACGACCAAGACCGGCAAACTTGGTCT 540
|||||
Db 9287 AAACCAATATGACCCCTATTCATTCACCAAGACGACCAAGACCGGCAAACTTGGTCT 9228
OY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACGTTGGAAGCAAGCTCAATAC 600
|||||
Db 9227 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACGTTGGAAGCAAGCTCAATAC 9168
OY 601 ACCAATTAATGATGACCTCAACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGA 660
|||||
Db 9167 ACCAATTAATGATGACCTCAACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGA 9108
OY 661 TAAATTAACCTCAAGGAAGTATATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACTCAC 720
|||||
Db 9107 TAAATTAACCTCAAGGAAGTATATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACTCAC 9048
OY 721 AACTGGATGATGATGCTTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGGGATAT 780
|||||
Db 9047 AACTGGATGATGATGCTTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGGGATAT 8988
OY 781 TTACAATACAGACGAATACGTTGCTATCCAGACGATGAATTCAGAGTGCCTTACAT 840
|||||
Db 8987 TTACAATACAGACGAATACGTTGCTATCCAGACGATGAATTCAGAGTGCCTTACAT 8928
OY 841 TGTGATGTTTCTAACGGTAAGTCAATGCCCCAGCTAGAGACGACGATCACTCAAGTAA 900
|||||
Db 8927 TGTGATGTTTCTAACGGTAAGTCAATGCCCCAGCTAGAGACGACGATCACTCAAGTAA 8868
OY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAATATGAA 960
|||||
Db 8867 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAATATGAA 8808
OY 961 ACCGATACAGACTATGCTGCTGCTGGAGTAGGCTGTCTAGATTCACACTGCTACTAT 1020
|||||
Db 8807 ACCGATACAGACTATGCTGCTGCTGGAGTAGGCTGTCTAGATTCACACTGCTACTAT 8748
OY 1021 CGTTACAGATGAGCCCTATTAACCTACCTGGAACAATTAACCTGTTTAACTGGGATAG 1080
|||||
Db 8747 CGTTACAGATGAGCCCTATTAACCTACCTGGAACAATTAACCTGTTTAACTGGGATAG 8688
OY 1081 GGGTACTTGGCAACATACCTGCAATAGGCCCGCAACATGCGCAAGAGTCCACAC 1140
|||||

QY 841 TGTGATGTTTCTAAGCGTAAAGTCATTGGCCAGCTAGAGGACGACGCGCATGTCAGTAA 900
|||||
Db 999 tgttgatgttcttaacggtaagtcacccagctagagcgacgacatccagtaaa 1058
QY 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAACAACCGGAGCTGGGATCACTATGAA 960
|||||
Db 1059 tgttctctcggaattacaacagtagaacaacacgagctgggtatcaactatgaa 1118
QY 961 ACCGATACAGACTATCTCTGCTTGAGTACGGTGTCTACATTCACCTGCTACTAT 1020
|||||
Db 1119 accgatcacgactatgtctctgctgagtagcgtatctacgtatcaactgtactat 1178
QY 1021 CGTTCAGATGAGCCCTATACCTGCGGACAAATACCTCTTTATACTGGATAG 1080
|||||
Db 1179 cgttcaagatgagcccatatacactcctggacaataaccctgttatactatgagtag 1238
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCACAACATCGGAAACGTCCAGC 1140
|||||
Db 1239 gggctactttggcaacatacacttgcaatcgccctgcaacaatcggaacgtccagc 1298
QY 1141 CGTGGAACTCTTAACAAGTCGAGCTCAACCGCGCAAGACTTCTTAATGCTTAGG 1200
|||||
Db 1299 cgttgaactctaaacaagctcgactcaacccgcaagacttccataatggtcgcg 1358
QY 1201 AATGACTACCCAGTATTTGCTAATGCAATGCAATTGCAATGCAACACGCAATCAGA 1260
|||||
Db 1359 aatcgactaccaagtatctcaactacatcaaatgcaatcctaagtaacaacacgactaga 1418
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGATGGCTGCTGCTTACGCTCCCTTTGCAATAG 1320
|||||
Db 1419 caaaaatattgagcaagtagtgaagaatgagtcgctgctttagcgtcgtccttgaatg 1478
QY 1321 TGGAACTTACTATTAACCAATGATATTCATTAAGTCTTATAGTAGGAGTGA 1380
|||||
Db 1479 tggaaacttactataacaactatgatalccataaagtcgtctttagtgaaggatgaaa 1538
QY 1381 AGAGTCTCAATGTGGAACTGCTGGCAAGAAACGACAGCCATATATGATGACCGA 1440
|||||
Db 1539 agagttctcaatgtcggaaactcgccatgaaagaaagacacgacataatgagacga 1598
QY 1441 CATGATGAAAAAGTCTTGAATGAGACTGACGAAATGCCATCTTGGCTTGGCTGCC 1500
|||||
Db 1599 catgatgaaaaacgctcttgactatgtaactgagcgaataatgctcttgcgtccccc 1658
QY 1501 TCAGGCTGTGTAACAGGAACCTCTTAATACTATACAGACGAGGAATGAAAAACATCAA 1560
|||||
Db 1659 tcaggtcgtgtaaaacaggaaccttaactatacagacgagaaattgaaaaccacataa 1718
QY 1561 GACCTCTCAATTTGTACACCTGATGAACTATTTGCTGGCTATACGGCTAAATTTCAAT 1620
|||||
Db 1719 gacctctcaatctgttagacctgtagaactatgtcgtgctatagcgttaataat 1778
QY 1621 GGCTGTATGAGAGGCTATTTCTTAACCTGTGACACCACTTGTAGCAATGGCTTACGGT 1680
|||||
Db 1779 ggcctgtatggaacggcctatcttaacggtctgacacacactgtgagcaatggccttaaggt 1838
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACCTGTCTGAAAGGAACATCCAGAGA 1740
|||||
Db 1839 cgcgtccaagtttaacgcctctatgactatctgtctgaaggaagcaatccagagga 1898
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGGAATTTGTAATTTAAATTTGCTCG 1800
|||||
Db 1899 ttggaataataccagagggctctacagaaatgagaaatcgttataaataatggtctcg 1958
QY 1801 TTCTACGTGGAACCTCCTCTCCACAACAACCCCATCAACTGAAGTTCAACCTCATC 1860
|||||
Db 1959 ttctacgttggaactcacctgtctcaacaacaccccaatcaactgaaagtccatc 2018
QY 1861 ATCAGATAGTTCAACTTACAGTCTAGCTCAACCATCCCAAGCACAATAATATAGTACGAC 1920
|||||
Db 2019 atcagatagttcaactcacagctcagctcaaccacacccaagcaacaataatagtaagac 2078

QY 1921 TACCAATCTTAACATAATAGCAACAATCAATACACCCCTGTATCAACAAATCAGAA 1980
|||||
Db 2079 taccaatccatacatataatagcaacaatacaataaaccctgatacaacaatacagaa 2138
QY 1981 TCCTCAACAGACACACCA 1999
|||||
Db 2139 tcccaaccagcacacca 2157

RESULT 7
US-09-107-433-1102
; Sequence 1102, Application US/09107433
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIA
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <unknown>
; OPERATING SYSTEM: <unknown>
; SOFTWARE: <unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2166 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...2166
; ; SEQUENCE DESCRIPTION: SEQ ID NO: 1102:
; US-09-107-433-1102

Query Match 99.4%; Score 1987.8; DB 15; Length 2166;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATATAGACATAATAATCACTATTCCTGACTTGGGTTCTAAGCCCGCGTCAA 60
|||||
Db 165 TAAATATAGACATAATAATCACTATTCCTGACTTGGGTTCTAAGCCCGCGTCAA 224
|||||
QY 61 TGCCCAAGCTAATGATTTTCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACCA 120
|||||

Db 225 TGCCAGAGCTAATGATATTCOCACAGATTTGGTTAAGGCAATGTTCTATCGAAGACA 284
QY 121 TCGCTTCTTGACACAGAGGGGATTTGATACATCCGATCCCTGGAGCTTCTTGCGCAA 180
Db 285 TCGCTTCTTGACACAGAGGGGATTTGATACATCCGATCCCTGGAGCTTCTTGCGCAA 344
QY 181 TCGCAAGAGCAATTCCTCCCAAGTGGATCACTCCACCAACAGTTGATTAAGTTGAC 240
Db 345 TCGCAAGAGCAATTCCTCCCAAGTGGATCACTCCACCAACAGTTGATTAAGTTGAC 404
QY 241 TTTACTTTTCACTTGCAGCTTCGACACAGACTATTTCTCGTAAGGCTCAGAGCTGGTT 300
Db 405 TTTACTTTTCACTTGCAGCTTCGACACAGACTATTTCTCGTAAGGCTCAGAGCTGGTT 464
QY 301 AGGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATTTGACCTATATAATAA 360
Db 465 AGGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATTTGACCTATATAATAA 524
QY 361 GGTCTACATGTCCTAATGGGAACTATGGAATGCAGACAGCTCAAAACTACTATGTAA 420
Db 525 GGTCTACATGTCCTAATGGGAACTATGGAATGCAGACAGCTCAAAACTACTATGTAA 584
QY 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCCCTTGCTGGTGAATGCTCAGCACC 480
Db 585 AGACCTCAATATTTAAGTTTACCTCAGTTAGCCCTTGCTGGTGAATGCTCAGCACC 644
QY 481 AAACCAATATGACCCCTATTCACATCAGAGAGGAGCCCAAGCCGCGAAACTTGCTT 540
Db 645 AAACCAATATGACCCCTATTCACATCAGAGAGGAGCCCAAGCCGCGAAACTTGCTT 704
QY 541 ATCTGAATATGAAAAATCAAGGCTACATCTCTGCTGAACAGATAGAAAAGCACTCAATAC 600
Db 705 ATCTGAATATGAAAAATCAAGGCTACATCTCTGCTGAACAGATAGAAAAGCACTCAATAC 764
QY 601 ACCAATATCTGAGAGCTACAAAGTCTCAATATAGCAAGTAATTAACCTCTTACATAGA 660
Db 765 ACCAATATCTGAGAGCTACAAAGTCTCAATATAGCAAGTAATTAACCTCTTACATAGA 824
QY 661 TTAATTCCTCAAGAGATCATCAATCAAGTTGAAGAAAGCAAGGCTATATACCTACTCAC 720
Db 825 TTAATTCCTCAAGAGATCATCAATCAAGTTGAAGAAAGCAAGGCTATATACCTACTCAC 884
QY 721 AACTGGATGATGATGTCTACACAATGTAGACAAGAGCTCAAAAACATCTGTGGATAT 780
Db 885 AACTGGATGATGATGTCTACACAATGTAGACAAGAGCTCAAAAACATCTGTGGATAT 944
QY 781 TTTACATACAGAGATAGCTTGCCTATCCAGACGATGAATTTGCAAGTCCCTCTACCAT 840
Db 945 TTTACATACAGAGATAGCTTGCCTATCCAGACGATGAATTTGCAAGTCCCTCTACCAT 1004
QY 841 TGTGTATGTTTCTAAGGATTAAGTCAATGTGCCAGCTAGAGACAGCCATAGTCAAGTAA 900
Db 1005 TGTGTATGTTTCTAAGGATTAAGTCAATGTGCCAGCTAGAGACAGCCATAGTCAAGTAA 1064
QY 901 TGTTCCTCTCGAATTAACCAAGCAGTAGAACAACCGCGAGCTGGGATCAACTATGAA 960
Db 1065 TGTTCCTCTCGAATTAACCAAGCAGTAGAACAACCGCGAGCTGGGATCAACTATGAA 1124
QY 961 ACCGATACAGACTATGCTCTCCCTGGAGTACGCTGTACGATTTCAACTCTCTACTAT 1020
Db 1125 ACCGATACAGACTATGCTCTCCCTGGAGTACGCTGTACGATTTCAACTCTCTACTAT 1184
QY 1021 CGTTACAGATGAGCCCTATTAATACCTGGAGCAAAATACCTGTTTATATACGCGATAG 1080
Db 1185 CGTTACAGATGAGCCCTATTAATACCTGGAGCAAAATACCTGTTTATATACGCGATAG 1244
QY 1081 GGGCTACTTTGGACACATCACTTGCATAATAGCCCTGCACAACAATCGGAAACGCTCCAGC 1140
Db 1245 GGGCTACTTTGGACACATCACTTGCATAATAGCCCTGCACAACAATCGGAAACGCTCCAGC 1304
QY 1141 CGTGGAAACTCTTAACAAGGTGGAGTCAACCGCGCAGAGACTTTCTTAATGTGTAGG 1200
Db 1305 CGTGGAAACTCTTAACAAGGTGGAGTCAACCGCGCAGAGACTTTCTTAATGTGTAGG 1364

QY 1201 AATCGACTACCCAGATATTCCTACTACTCAAAATGCCATTTCAAGTAACACACCAATCAGA 1260
Db 1365 AATCGACTACCCAGATATTCCTACTACTCAAAATGCCATTTCAAGTAACACACCAATCAGA 1424
QY 1261 CAAAAAATATGAGCAAGTAGTAAGAAAGATGCTGCTCTTACAGCTGCTCTTCCAAATGG 1320
Db 1425 CAAAAAATATGAGCAAGTAGTAAGAAAGATGCTGCTCTTACAGCTGCTCTTCCAAATGG 1484
QY 1321 TGGAACTTACTATAAACCATGATATCCATTAAGTCGTCTTATAGATGGGAGTGAATA 1380
Db 1485 TGGAACTTACTATAAACCATGATATCCATTAAGTCGTCTTATAGATGGGAGTGAATA 1544
QY 1381 AGAGTTCTCTAATGTGGAAGCTGTCGCATGAGAAGAAACAGACAGCTATATGATGACCA 1440
Db 1545 AGAGTTCTCTAATGTGGAAGCTGTCGCATGAGAAGAAACAGACAGCTATATGATGACCA 1604
QY 1441 CATGATGAAAAACAGCTCTTACTATATGGAACCTGGAACGAATGCTATCTTGCTGCC 1500
Db 1605 CATGATGAAAAACAGCTCTTACTATATGGAACCTGGAACGAATGCTATCTTGCTGCC 1664
QY 1501 TCAAGCTGGTAAACAGGAACCTCTAATATACAGACAGAGGAATTTGAACCAATCA 1560
Db 1665 TCAAGCTGGTAAACAGGAACCTCTAATATACAGACAGAGGAATTTGAACCAATCA 1724
QY 1561 GACCTCTCAATTTGTAGACACCTGATGAACTATTTGCTGGCTATACGGTAATATTCAT 1620
Db 1725 GACCTCTCAATTTGTAGACACCTGATGAACTATTTGCTGGCTATACGGTAATATTCAT 1784
QY 1621 GAGCTATATGACAGGCTATTTCTAACCGCTTGACACCACTTTGAGCAATGCTTACGGT 1680
Db 1785 GAGCTATATGACAGGCTATTTCTAACCGCTTGACACCACTTTGAGCAATGCTTACGGT 1844
QY 1681 CGCTGCCAAAGTTTACCGCTATGAGTACCTGCTGCTGAAGGAAGCAATCCAGACGA 1740
Db 1845 CGCTGCCAAAGTTTACCGCTATGAGTACCTGCTGCTGAAGGAAGCAATCCAGACGA 1904
QY 1741 TTGGAATATACAGAGGAGGCTCTACAGAAATGAGAAATTCGATTTAAAAATGGTCTCG 1800
Db 1905 TTGGAATATACAGAGGAGGCTCTACAGAAATGAGAAATTCGATTTAAAAATGGTCTCG 1964
QY 1801 TTCTACGTGGAAGTCACTGCTGCTCACACACACCCCATCACTGAAATGTTCAAGCTCATC 1860
Db 1965 TTCTACGTGGAAGTCACTGCTGCTCACACACACCCCATCACTGAAATGTTCAAGCTCATC 2024
QY 1861 ATCAGATAGTTCAAGTCTCAGCTAGCTCAACCTCCAGACCAATATATGATAGCAG 1920
Db 2025 ATCAGATAGTTCAAGTCTCAGCTAGCTCAACCTCCAGACCAATATATGATAGCAG 2084
QY 1921 TACCAATCTTAACAAATATACGACAAATCAATTAACACCCCTGATCAACAATATGAGAA 1980
Db 2085 TACCAATCTTAACAAATATACGACAAATCAATTAACACCCCTGATCAACAATATGAGAA 2144
QY 1981 TCCTCAACGACACAACA 1999
Db 2145 TCCTCAACGACACAACA 2163

RESULT 8
PCT-US02-03987-9325
; Sequence 9325, Application PC/TUS0203987
; GENERAL INFORMATION:
; Applicant: Elitza Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the target of a Compound which inhibit
; FILE REFERENCE: ELITRA-028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; PRIOR FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9325

LENGTH: 2160
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(2160)
PCT-US02-03987-9325

Query Match 99.1%; Score 1981.4; DB 1; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTCGACATTAATCAACTCATGTGCTGGGTTCGAGCGCGCTCA 60
Db 159 TAAATCTCGACATTAATCAACTCATGTGCTGGGTTCGAGCGCGCTCA 218
QY 61 TCCCAAGCTAATGATATCCACAGATTGGTTAGGCAATCGTTCTATGAGACCA 120
Db 219 TCCCAAGCTAATGATATCCACAGATTGGTTAGGCAATCGTTCTATGAGACCA 278
QY 121 TCGCTTCTCGACACACAGGGGATGATACATCGTATCTGGAGCTTTCTGCGCA 180
Db 279 TCGCTTCTCGACACACAGGGGATGATACATCGTATCTGGAGCTTTCTGCGCA 338
QY 181 TCGCAAGCAATTCCTCCAGGTGATCACTCTCACCCCAAGTGTGATTAAGTTGAC 240
Db 339 TCGCAAGCAATTCCTCCAGGTGATCACTCTCACCCCAAGTGTGATTAAGTTGAC 398
QY 241 TTACTTTTCACTGACTTCGACAGACTATTTCTCGTAAGGCTCAGAAAGCTTGTT 300
Db 399 TTACTTTTCACTGACTTCGACAGACTATTTCTCGTAAGGCTCAGAAAGCTTGTT 458
QY 301 ACGATTCAGTTAGCAAAAAGCAACAAGCAAGAATCTTGACTTATTAATAA 360
Db 459 ACGATTCAGTTAGCAAAAAGCAACAAGCAAGAATCTTGACTTATTAATAA 518
QY 361 GGTCTCATGTCAATGGAACATGATGACAGACAGACGTCAAATACATAGTGA 420
Db 519 GGTCTCATGTCAATGGAACATGATGACAGACAGACGTCAAATACATAGTGA 578
QY 421 AACCTCAATTAATTAAGTTTACCTCAGTTAGGCTTGCTGGATGCTTCAGCAC 480
Db 579 AACCTCAATTAATTAAGTTTACCTCAGTTAGGCTTGCTGGATGCTTCAGCAC 638
QY 481 AAACCAATGACCCCTATTCACATCCAGAACGACCCCAAGCCGCGAACTTGCTT 540
Db 639 AAACCAATGACCCCTATTCACATCCAGAACGACCCCAAGCCGCGAACTTGCTT 698
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGTGAACGATAGAAAGCATCAATAC 600
Db 699 ATCTGAAATGAAAAATCAAGGCTACATCTCTGTGAACGATAGAAAGCATCAATAC 758
QY 601 ACCAATTAATGATGACATCAAAAGTCTCAAAATCAGCAAGTAATTAATCTGCTACATGA 660
Db 759 ACCAATTAATGATGACATCAAAAGTCTCAAAATCAGCAAGTAATTAATCTGCTACATGA 818
QY 661 TAATTACGTCAGAGATCATCAATCAAGTTGAAGAAGAAACGCTATTAACCTACTAC 720
Db 819 TAATTACGTCAGAGATCATCAATCAAGTTGAAGAAGAAACGCTATTAACCTACTAC 878
QY 721 AACGGGATGATGATGATCAAAATGTAGACCAAGAAAGCTCAAAACATCTGGGATAT 780
Db 879 AACGGGATGATGATGATCAAAATGTAGACCAAGAAAGCTCAAAACATCTGGGATAT 938
QY 781 TTACATTAACAGAGATGATGCTATCCAGAGATGAATTCAGATGCTGCTCTACAT 840
Db 939 TTACATTAACAGAGATGATGCTATCCAGAGATGAATTCAGATGCTGCTCTACAT 998
QY 841 TGTGATGTTTCTAACGGTAAAGTCAATGGCCGCTAGAGACAGCCATCAAGTCAAGTAA 900
Db 999 TGTGATGTTTCTAACGGTAAAGTCAATGGCCGCTAGAGACAGCCATCAAGTCAAGTAA 1058

QY 901 TGTTCCTTCGGAAATTAACCAAGCAGTAGAAACAAACCGCAGCTGGGATCAATATGA 960
Db 1059 TGTTCCTTCGGAAATTAACCAAGCAGTAGAAACAAACCGCAGCTGGGATCAATATGA 1118
QY 961 ACCGATACACACTATGCTCTGCTGCTGAGTACGCTGTCTACAGATTCATCTACTAT 1020
Db 1119 ACCGATACACACTATGCTCTGCTGCTGAGTACGCTGTCTACAGATTCATCTACTAT 1178
QY 1021 CGTTACGATGAGACCCCTATTAATCACTACCTGGGACAAATTAATCTGTTATTAACGGGATAG 1080
Db 1179 CGTTACGATGAGACCCCTATTAATCACTACCTGGGACAAATTAATCTGTTATTAACGGGATAG 1238
QY 1081 GGGCTACTTGGCAACATCACTTGAATATGAGCTGCAACATCGCAAAACGCTCCACG 1140
Db 1239 GGGCTACTTGGCAACATCACTTGAATATGAGCTGCAACATCGCAAAACGCTCCACG 1298
QY 1141 CGTGAACCTTAACCAAGGTGAGCTCAACCGCGCAAGACTTTCTTAATGCTCTAG 1200
Db 1299 CGTGAACCTTAACCAAGGTGAGCTCAACCGCGCAAGACTTTCTTAATGCTCTAG 1358
QY 1201 AATCGACTACCCCAAGTATTCATCAATGCAATTCATTAAGTAACAAACGCAATCAGA 1260
Db 1359 AATCGACTACCCCAAGTATTCATCAATGCAATTCATTAAGTAACAAACGCAATCAGA 1418
QY 1261 CAAAAATATGAGCAAGTAGTGAAGAATGCTGCTGCTTACGCTGCTTCAATAG 1320
Db 1419 CAAAAATATGAGCAAGTAGTGAAGAATGCTGCTGCTTACGCTGCTTCAATAG 1478
QY 1321 TGGAACTTACTATTAACCAATGATATTCATTAAGTGAAGTGAAGTGAAGTGA 1380
Db 1479 TGGAACTTACTATTAACCAATGATATTCATTAAGTGAAGTGAAGTGAAGTGAAGTGA 1538
QY 1381 AGAGTCTCTAATGCTGGAACCTGCTGCCATGAAAGAAAGACGCTTATATGATGACCA 1440
Db 1539 AGAGTCTCTAATGCTGGAACCTGCTGCCATGAAAGAAAGACGCTTATATGATGACCA 1598
QY 1441 CATGATGAAGAAAGCTTATGACTATGAACTGACCAAAATGCTTATCTGCTGGCTCC 1500
Db 1599 CATGATGAAGAAAGCTTATGACTATGAACTGACCAAAATGCTTATCTGCTGGCTCC 1658
QY 1501 TCAAGCTGTTAAACAGGAACCTCTAATATACAGACGAGGAATTAAGAAACCAATCA 1560
Db 1659 TCAAGCTGTTAAACAGGAACCTCTAATATACAGACGAGGAATTAAGAAACCAATCA 1718
QY 1561 GACCTCTCAATTTGTAGCACCTGATGACTATTTGCTGGCTATACGCTTAATATTCAT 1620
Db 1719 GACCTCTCAATTTGTAGCACCTGATGACTATTTGCTGGCTATACGCTTAATATTCAT 1778
QY 1621 GGCTGATGAGACGCTATCTAAGCGTCTGACACCACTGTGAGCAATGAGCTTACGGT 1680
Db 1779 GGCTGATGAGACGCTATCTAAGCGTCTGACACCACTGTGAGCAATGAGCTTACGGT 1838
QY 1681 CGCTGCCAAAGTTTACCGCTATGATGACCTACCTGCTGGAAGAACCAATCCAGAAGA 1740
Db 1839 CGCTGCCAAAGTTTACCGCTATGATGACCTACCTGCTGGAAGAACCAATCCAGAAGA 1898
QY 1741 TTGGAATATACGAGAGGCTCTACAGAAATGGAATTCGTTTAAAAATGCTGCTG 1800
Db 1899 TTGGAATATACGAGAGGCTCTACAGAAATGGAATTCGTTTAAAAATGCTGCTG 1958
QY 1801 TTCTAGTGAACCTCACTGCTCCACAAACCAACCCCATCAACATGAAGTTCAAGCTCAT 1860
Db 1959 TTCTAGTGAACCTCACTGCTCCACAAACCAACCCCATCAACATGAAGTTCAAGCTCAT 2018
QY 1861 ATCAGATAGTTTCAACTTCACAGTCTAGCTCAACACTCCAGACACAATATATAGTACAG 1920
Db 2019 ATCAGATAGTTTCAACTTCACAGTCTAGCTCAACACTCCAGACACAATATATAGTACAG 2078
QY 1921 TACCAATCTTAACAAATATATACGCAACATCAATTAACACCCCTGATCAACAAAATCAGAA 1980
Db 2079 TACCAATCTTAACAAATATATACGCAACATCAATTAACACCCCTGATCAACAAAATCAGAA 2138
QY 1981 TTCTCAACGACACAACCA 1999

Db 2139 tccctaccagcaacacca 2157

RESULT 9

us-09-815-242-9325

Sequence 9325, Application US/09815242

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zykkind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011a

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9325

LENGTH: 2160

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(2160)

US-09-815-242-9325

Query Match 99.18; Score 1981.4; DB 31; Length 2160;
Best Local Similarity 99.48; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 TAAATTCACACATAAATAATCAATCATGCTGACTTGGTTCTGAACGGCGGCTCAA 60
159 taataatctagacataataataatcaatcatgctgacttgggttctgaagcgcggtcaa 218
61 TGCCCAAGCTAATGATATTCACAGATTTGTTAAGGCAATCGTTTCTATCGAAGACCA 120
219 tgcacaagctaataatgataatctccacagatttggtaaggcaatcgttctatcgaagacca 278
121 TCGCTTCTTCACACACAGGGGATTTGATACCATCGTATCTGGAGCTTTCTTCGCGAA 180
279 tcgcttcttcgacacaggggagatgatacaccatcgatctcgtggagcttcttcgcaaa 338
181 TCTGCAAGCAATTCCTCCCAAGGCGATCACTCCACCAACAGTTGATTAAAGTTGAC 240
339 tctgcaagcaatctccccaaggtgagctccacccaacagttgattaaagttgac 398
241 TTACTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTGGTT 300
399 ttactttcaacttcgacttcgacacagactatttctcgtaaagctcaggaagcttggt 458
301 AGCGATTAGTTAGAACAAAAGCACCAACAGAAATCTTGACCTATATTAATAATA 360

Db 459 agcgatcagltagagacaaaagcaaccagaagaatcttgacctactatataataaa 518
361 GGTCTACATGCTTAATGGGAACTATGATGCAGACACAGCTCAAAACTACTATGTAA 420
519 ggtctacatgcttaatgggaactatagatgacagcagctcaaaactactatagttaa 578
421 AGACCTCAATATATTTAAGTTTACTCTGACTTACCTTGTGCTGGATATGCTCAGGACC 480
579 agacctcaaatatattaaagtttaacctcagttagcccttgctgctgaaagctcagccacc 638
481 AAACCAATATGACCCCTTTTCACATCCGAGAGCCCAAGCCGCGCAAACTTGTGTT 540
639 aaaccaataatgaccccttcaactcagaagcagcccaagcgcgaagctcgtgctc 698
541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATATAC 600
699 atctgaatgaaaaatcaaggctacatctctcgtgaacagctatgagaagaagtaataaac 758
601 ACCAATTACTGATGACTTACAAAGTCTCAATCAGCAAGTAAATTAACCTGCTTACATGGA 660
759 accaattactgatgacttacaagctcacaagctcaaatcagaagtaataacctgttacaatgga 818
661 TAAATTACCTCAAGGAGTCATCAATCAAGTTGAGAAAGAAAGGCTTAACCTTACATCAC 720
819 taattacctaaggaagctcaatcaatcaagttgaagaagaagaagctataacctaactaac 878
721 AACTGGATGATGCTCTACACAAATGTAGACCAGAAAGCTCAAAACATCTGTGGGATAT 780
879 aactggatgatgctctacacaaatgtagaccagaaagctcaaaaacatctgtggatatt 938
781 TTACAAATCAGACGAAATACGTTGCTTATCCAGACGATGAATTTGCAAGTCTTCTACCAT 840
939 ttacaaatcagacgaaatacgttgcttattccagacgatatgaaagctccttaccat 998
841 TGTGTATGTTCTTAAGGTTAAAGTCATGCCCAGCTAGACAGCGCTTCAGTCAAGTAA 900
999 tgtgtatgttcttaaggtttaagtcattgcccagctaggaagcaagcctcagtaagtaa 1058
901 TGTTCCTTCGCAATTAACCAAGCAGTAGAAACAAACCGCAGCTGGGGATCACTATGAA 960
1059 tgttccttcgcaattaaccaaagcagtagaagaacaacgcgagctgggataactaactaga 1118
961 ACCGATCAGACGATATGCTCTGCTTGGAGTAGGGTGTCTACGATTTCACTGCTACTAT 1020
1119 accgatcagacgatatgctctgcttggagtagagtgctacagtgtaactgaactat 1178
1021 CGTTCAGATGATGCGCTTAACTACCTACCCGAGCAAAATCTGTTTAACTGGGATAG 1080
1179 cgttcagatgatgcgctttaaactacccctggagcaaaatctgttttaaactgggataag 1238
1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTCGCAAAATCGCGAAGCGTCCAGC 1140
1239 gggctactttggcaaacatcaccttgcaatatcgccctcgaaacaatcgcgaagcgtccagc 1298
1141 CGTGGAAACTCTTAACAAGGTCGACTACACCGCGCAAGCTTCTTAATGCTTAGG 1200
1299 cgtggaaactctttaacaaggctacacccgcaacgcgcgaagcttcttaaatgcttagg 1358
1201 AATGCACTACCCAGTATTAATCACTACTCAAAATGCCATTTCAAGTAAACAAACGGAATGGA 1260
1359 aatgcactacccagtatattcaactcaaaatgccattcaagtaaacacaacggaatcaga 1418
1261 CAAAAAATATGAGACAAGTATGAAAAAGATGGCTGCTTACCTGCTTGGCAAAATG 1320
1419 caaaaaatataatgagacaagtatgaaaaagatgctgctgcttaagctccttgcaaatg 1478
1321 TGGAACTTACTATAACCAATGTATATCCATTAAGCTCTTTAGTGTGGAGTGAATA 1380
1479 tggaaacttactataaaccaatgtatatccataaagctgctttagtgaaggagtgaaaa 1538
1381 AGAGTTCTTAATGTGGAACATCGTGCATATGAGAAAGCAAGCAGCTATATGATGACCGA 1440
1539 agagttcttaatgtggaacatcgatgcataatgagaaagcaagcagctatatatgaccca 1598

```
QY 1441 CATGATGAAAGAGTCTTACTTATGAGAGTGGAGAAATGCTATCTTGCTGGCTCCC 1500
    |||
Db 1599 catgatgaaagagctcttacttatagtatgaaacttgagaaatgctatcttgctgctccc 1658
QY 1501 TCAGGCTGGTAAAGAGAAAGAACTCTACTATACAGAGAGAAATTAAGAAACCATCAAA 1560
    |||
Db 1659 tcaggctggtaaagaaagaaactctaacatatacagcaggaattgaaacacatcaa 1718
QY 1561 GACCTCTCAATTTGTAGCAGCTGATGAACTATTTGCTGCTATACGCGTAATATTCAAT 1620
    |||
Db 1719 gacctctcaattttagcagccgtagtaactatttgcctgctatacgcgtlaaatatcaat 1778
QY 1621 GCCTGATGAGAGGCTATTTTACCGCTGACACCACTTGTAAGCAATGGGCTTTACGCT 1680
    |||
Db 1779 ggcctgtagagagctatcttaaccgctctgacacacatctgtagcaatgaccttaagct 1838
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGAGCTACCTGCTGTAAGAGCAATCCAGAGAA 1740
    |||
Db 1839 cgctgccaaagttaaccgctctatgtagtacctgctctgtagaggaatccagagga 1898
QY 1741 TTGGAATATACAGAGAGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTCTCG 1800
    |||
Db 1899 ttggaatatagcagagggctctacagaaatgagaaatcgtaatttaaaatggtctcg 1958
QY 1801 TTTCTAGTGAAGTCACTGCTGCTCCACAAACACCCCATGCACTGAAGTTCAAGCTCATC 1860
    |||
Db 1959 ttctaagtagagctcaccgctctccacaaaccccatcaactgaagttcaagctcatc 2018
QY 1861 ATCAGATAGTTCACACTTCAAGTCTAGTCAACACTCCAGACAAATTAATAGTAGCAGC 1920
    |||
Db 2019 atcagatagttcaacttcaactcagctcagctcaacacactccaagcacaataatagtcagc 2078
QY 1921 TACCAATCTTCAATATATAGCGAACAAATCAACACCCCTGATCAACAAATATCAGAA 1980
    |||
Db 2079 taaccaatcttaacaataatagcaacaatcaataacacccctgatacaazaatcagaa 2138
QY 1981 TCCTCAACAGACAGACACCA 1999
    |||
Db 2139 tccctcaacagacacacca 2157

RESULT 10
US-10-072-851-9325
; Sequence 9325, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jhang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
US-10-072-851-9325

Query Match          99.1%; Score 1981.4; DB 37; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATAAATCACTATGCTGACTGGTTCGTAAGCCGCGCTCAA 60
    |||
Db 159 taaatctacgacaaataaaatcaactatgctgactgggttcgtaagccgcgctcaa 218
QY 61 TGGCCAGCTAATGATATTCCACAGATTGGTTAAGGCAATCGTTTATGCAACCA 120
    |||
Db 219 tggccagctaatgatatccacagatttggtaaggaatgcttctatcgaagacca 278
QY 121 TGGCTCTTGACACACAGGGGGATTGATACATCCGTAATCGTGGAGCTTTCTGGCAA 180
    |||
Db 279 tggctcttgacacacaggggattgataccatcgatccgtagggagcttcttgccaa 338
QY 181 TCTGCAAGCAATTCCTCCAGAGGTGATCACTCTACCCCAACAGTTGATTAACTGAC 240
    |||
Db 339 tctgcaagcaattcctccagaggtgatacagctcactcaacagttgataagttgac 398
QY 241 TTTACTTTTCACTTGACCTTCCGACGACATTTTCTGTAAGCTCAGAGAGCTTGT 300
    |||
Db 399 ttactttcaacttgcgacttcgaccagactattctcgtlaagctcgaagacttgct 458
QY 301 AGCGATTCACTTGAACAAAGAACCAACCAAGCAAGAAATCTTGACTATTAATAA 360
    |||
Db 459 agcgattcagttagaacaaagaaacaaagaaatcttgaactactatataataa 518
QY 361 GGTCTACATGTCTAATGGAACATGTAATGCAGACAGCAAGCTCAAACTACTATGTA 420
    |||
Db 519 ggtctacatgtctaatgggaactatggaatgacgaagaagctcaaaactactatgtaa 578
QY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGGCTTGCTGTAAGTCCCTCAGCACC 480
    |||
Db 579 agacctcaatattaagtttacctcagttagcttgcgtgtaagtccctcagccacc 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGACAGCCCAAGCCGCAAACTTGCTCT 540
    |||
Db 639 aaaccaatatagacccctataccaatcagaagaagcccaagccggaacttgctt 698
QY 541 AACTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
    |||
Db 699 aactgaaatgaaaaatcaaggtcacaatctctgctgaacagttatgaaaagccaatc 758
QY 601 ACCAATTACTGATGACTACAAAGTCTCAATCAGCAAGTAAATTACCTGCTTACATGGA 660
    |||
Db 759 accaattactgatgactacaagctcaaatcgaatcgaagaaagaaagcttaccatgga 818
QY 661 TAATTACCTCAAGAGAGTATCAATCAAGTTGAAAGAAAGAGGTATTAACCTATCTAC 720
    |||
Db 819 taattacctcaagagagtatcaatcaagttgaaagaaagaggtataaactactcac 878
QY 721 AACTGGATGATGCTGTACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGAGTAT 780
    |||
Db 879 aactggatgatgctgtacaaatgtagaccagaagctcaaaaacatctgtggagat 938
QY 781 TTACAATACAGAGCAATAGCTTGCTTACAGAGATGATGCAAGTGCCTTACCAT 840
    |||
Db 939 ttacaatacagagcaatagcttgcttaccagagtgatattgcaagctccttaccat 998
QY 841 TGTGATGTTTCTAGCGGTAAAGTCTTGGCCAGCTAGAGACAGCGCATCAGTCAAGTAA 900
    |||
Db 999 tgtgatgtttctagcggtaaagcttggccagctagagacagcgcacatcagttcaagtaa 1058
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGAGTCAACTATGAA 960
    |||
Db 1059 tgttccttcggaattaaaccaagcagtagaacaacacccgagctgggagctcaactatgaa 1118
```


301 AGCGATTCAGTTAGAACAAAAACCAACGACGAAATCTTGACCTACTATATAATA 360
|||||
9418 AGCGATTCAGTTAGAACAAAAACCAACGACGAAATCTTGACCTACTATATAATA 9359
361 GGTCTACATGCTTAATGGGACTATGGAATGAGACAGCAGCTCAAACTACTATGTA 420
|||||
9358 GGTCTACATGCTTAATGGGACTATGGAATGAGACAGCAGCTCAAACTACTATGTA 9299
421 AGACCTCATATATTTAAGTTTACCTCAGTATACCTTGCTGGCTGGAAATGCTCAGCACC 480
9298 AACCTCATATATTTAAGTTTACCTCAGTATACCTTGCTGGCTGGAAATGCTCAGCACC 9239
481 AACCAATATGACCCCTATTTCATCAGACAGAGCCCAAGACCGCCCAATGCTGCT 540
9238 AACCAATATGACCCCTATTTCATCAGACAGAGCCCAAGACCGCCCAATGCTGCT 9179
541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGTATGAGAAAGCATCAATAC 600
9178 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGTATGAGAAAGCATCAATAC 9119
601 ACCAATCTGATGAGACATCAAAAGTCTCAATCAGCAGTAAATTCCTGCTTACATGGA 660
9118 ACCAATCTGATGAGACATCAAAAGTCTCAATCAGCAGTAAATTCCTGCTTACATGGA 9059
661 TAATTAACCTCAAGGAGTCAATCAATGTAAGAGAAAGAGGCTATTAACCTACTCAC 720
9058 TAATTAACCTCAAGGAGTCAATCAATGTAAGAGAAAGAGGCTATTAACCTACTCAC 8999
721 AACTGGAGTGAATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
8998 AACTGGAGTGAATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 8939
781 TTACATATACAGACGATAGCTGCTATCCAGACGATGAATTCGCAAGTGGCTTACCAT 840
8938 TTACATATACAGACGATAGCTGCTATCCAGACGATGAATTCGCAAGTGGCTTACCAT 8879
841 TCTTATGATGTTTCTAAGGTAAGTCAATGCCAGCTAGAGAGCAGCATCAAGTAA 900
8878 TCTTATGATGTTTCTAAGGTAAGTCAATGCCAGCTAGAGAGCAGCATCAAGTAA 8819
901 TCTTATGATGTTTCTAAGGTAAGTCAATGCCAGCTAGAGAGCAGCATCAAGTAA 960
8818 TCTTATGATGTTTCTAAGGTAAGTCAATGCCAGCTAGAGAGCAGCATCAAGTAA 8759
961 ACCGATACAGACTATGCTGCTGCTGAGTACGCTGCTACGATTCACCTGCTACTAT 1020
8758 ACCGATACAGACTATGCTGCTGCTGAGTACGCTGCTACGATTCACCTGCTACTAT 8699
1021 CGTTACGATGAGCCCTATTAACCTGAGCAAAATATCTCTGTTTAACTGGGATAG 1080
8698 CGTTACGATGAGCCCTATTAACCTGAGCAAAATATCTCTGTTTAACTGGGATAG 8639
1081 GGGCTACTTTGGCAACATCACTGCAATATGACCCCTGCAACAATGCGAAAGCTCCACG 1140
8638 GGGCTACTTTGGCAACATCACTGCAATATGACCCCTGCAACAATGCGAAAGCTCCACG 8579
1141 CGTGAACCTTAACACAGGTGAGCTAACCGCGCAAGACTTTCCTTAATAGTCTAGG 1200
8578 CGTGAACCTTAACACAGGTGAGCTAACCGCGCAAGACTTTCCTTAATAGTCTAGG 8519
1201 AATGCACTACCAAGATTAATCACTCAATGCAATTCGAATTAACACAGCAATCAGA 1260
8518 AATGCACTACCAAGATTAATCACTCAATGCAATTCGAATTAACACAGCAATCAGA 8459
1261 CAAAAAATATGAGCAAGTATGTAAGATGCTGCTGCTACGCTGCTTTGCAATAG 1320
8458 CAAAAAATATGAGCAAGTATGTAAGATGCTGCTGCTACGCTGCTTTGCAATAG 8399
1321 TGGAACTTACTATAACCAATGATATATCATATAAGTCTGCTTATAGTGGAGTGA 1380
8398 TGGAACTTACTATAACCAATGATATATCATATAAGTCTGCTTATAGTGGAGTGA 8339
1381 AGAGTTCTCTAATGTGCAACTGTGCGCATGGAAGGAAAGCAGACCTATATGATGACGA 1440

|||||
8338 AGAGTTCTCTAATGTGCAACTGTGCGCATGGAAGGAAAGCAGCCTATATGATGACGA 8279
1441 CATGATGAAACAGCTTACTTATGGAACCTGAGGAAATGCTTATCTTGCTGCTCC 1500
8278 CATGATGAAACAGCTTACTTATGGAACCTGAGGAAATGCTTATCTTGCTGCTCC 8219
1501 TCAGGCTGTAAACAGGACCTTAACTATACAGACGAGGAAATGGAACACATCA 1560
8218 TCAGGCTGTAAACAGGACCTTAACTATACAGACGAGGAAATGGAACACATCA 8159
1561 GACCTCTCAATTTAGACACCTGATGAATATTTGCTGGCTTACGCTAAATTTCAAT 1620
8158 GACCTCTCAATTTAGACACCTGATGAATATTTGCTGGCTTACGCTAAATTTCAAT 8099
1621 GCTGTATGAGACAGCTTATTAACCGTGTGACACCACTTGTAGCAATGSCCTTACGCT 1680
8098 GCTGTATGAGACAGCTTATTAACCGTGTGACACCACTTGTAGCAATGSCCTTACGCT 8039
1681 CGCTGCCAAAGTTTACCGCTTATGATGACCTACCTGTCTGAAGGAAAGCAATCCAGAGA 1740
8038 CGCTGCCAAAGTTTACCGCTTATGATGACCTACCTGTCTGAAGGAAAGCAATCCAGAGA 7979
1741 TTGGAATATACAGAGGAGGCTCTACAGAAATGAGAAATCGATTTAAATAGTGTCTG 1800
7978 TTGGAATATACAGAGGAGGCTCTACAGAAATGAGAAATCGATTTAAATAGTGTCTG 7919
1801 TTCTACGTGGAACCTACCTGCTCCACACACACCCCATCACTGAAAGTTCAGCTCATC 1860
7918 TTCTACGTGGAACCTACCTGCTCCACACACACCCCATCACTGAAAGTTCAGCTCATC 7859
1861 ATCAGATAGTTAAGCTTCACTGACCTGACCTCAACCACTCCACCAAAATATAGTACGAC 1920
7858 ATCAGATAGTTAAGCTTCACTGACCTGACCTCAACCACTCCACCAAAATATAGTACGAC 7799
1921 TACCAATCCTAACAATATATACGCAACCAATCAATCAACCCCTGATTCACAAATAGTACGA 1980
7798 TACCAATCCTAACAATATATAGCACAACCAATCAATCAACCCCTGATTCACAAATAGTACGA 7739
1981 TCCTCAACGACACACCA 1999
7738 TCCTCAACGACACACCA 7720

RESULT 12
US-60-068-175-596/c
; Sequence 596, Application US/60068175
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; APPLICANT: Floney, Gregory L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
; NUMBER OF SEQUENCES: 1175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068.175
; FILING DATE: HERewith
; CLASSIFICATION:

```
ATTORNEY/AGENT INFORMATION:
:
: NAME: CERRONE, MICHAEL C.
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PM-0009-2 P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-416
: INFORMATION FOR SEQ ID NO: 596:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9845 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: genomic DNA
: IMMEDIATE SOURCE:
: CLONE: SPN2C598
: US-60-068-175-596

Query Match      81.6%; Score 1631.2; DB 45; Length 9845;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1776; Conservative 0; Mismatches 223; Indels 1; Gaps 1;

OY 1 TAAATCTACGCAATAAATCACTGCTGCTGCTGGTTCTGAACCGCGCTCAA 60
Db 9734 TAAATCTACGCAATAAATCACTGCTGCTGCTGGTTCTGAACCGCGCTCAA 9675

OY 61 TGCCCAAGCTATGATATTTCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACCA 120
Db 9674 TGCCCAAGCTATGATATTTCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACCA 9615

OY 121 TCGCTTCTTCGACACAGAGGGGATGATACATCCGTAATCTGGGAGCTTTCTTGCGCAA 180
Db 9614 TCGCTTCTTCGACACAGAGGGGATGATACATCCGTAATCTGGGAGCTTTCTTGCGCAA 9555

OY 181 TCTGCAAGCAATCCCTCCAGGTGATGCACTCTCAACCAAGTGTATTAATGTAC 240
Db 9554 TCTGCAAGCAATCCCTCCAGGTGATGCACTCTCAACCAAGTGTATTAATGTAC 9495

OY 241 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGAAAGTTGGTT 300
Db 9494 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGAAAGTTGGTT 9435

OY 301 AGCGATTGATAGAACAAAAAGCAACCAAGCAAGAAATCTTGACTACTATATAATTA 360
Db 9434 AGCGATTGATAGAACAAAAAGCAACCAAGCAAGAAATCTTGACTACTATATAATTA 9375

OY 361 GGTCTACATGCTATGGAAGCTATGGAATGAGACAGACAGCTCAAACTACTATGTTAA 420
Db 9374 GGTCTACATGCTATGGAAGCTATGGAATGAGACAGACAGCTCAAACTACTATGTTAA 9315

OY 421 AGACCTCAATATTTAAGTTTACCTCACTAGCTTGTGCTGGATGATGCTCAGAGCAC 480
Db 9314 AGACCTCAATATTTAAGTTTACCTCACTAGCTTGTGCTGGATGATGCTCAGAGCAC 9255

OY 481 AAACCAATATACCCCTTATTCATCATCGAAGACAGCCCAAGACGCGCGAAATCTGCTT 540
Db 9254 AAACCAATATACCCCTTATTCATCATCGAAGACAGCCCAAGACGCGCGAAATCTGCTT 9195

OY 541 ATCTGAATGAAAAATCAAGGCTACTCTGCTGGAACAGTATGGAAGAGTCAATAC 600
Db 9194 ATCTGAATGAAAAATCAAGGCTACTCTGCTGGAACAGTATGGAAGAGTCAATAC 9135

OY 601 ACCAATTAAGTGAAGTCTCAATTCAGCAAGTATTAACCTT- GCTTACATAGG 659
Db 9134 ACCAATTAAGTGAAGTCTCAATTCAGCAAGTATTAACCTT- GCTTACATAGG 9075

OY 660 ATATATTACCTTAAGAGACTCATCAATCAAGTTGAAGAAAGAGGCTATTAACCTACTCA 719
Db 9074 ATATATTACCTTAAGAGACTCATCAATCAAGTTGAAGAAAGAGGCTATTAACCTACTCA 9015

OY 720 CAACGGGATGATGTCTACCAATGTAGCAAGAAAGCTCAAAACATCTGTTGGGATA 779
Db 720 CAACGGGATGATGTCTACCAATGTAGCAAGAAAGCTCAAAACATCTGTTGGGATA 779

OY 9014 CTACTGGATGATGTTTACCAAAATGTAGACCAAGAAAGCTCAAAACATCTGTTGGGATA 8955
OY 780 TTTACAATACAGACGAATACGTTGGCTATCCAGACGATGAATTGGCAAGCTTCCTACCA 839
Db 8954 TCTCAACCTCCGATCAATACGTTGGCTATCCAGACGATGAATTGGCAAGCTTCCTACCA 8895

OY 840 TTTGTTGATGTTTCTAACGCTTAAGTCAATGCTGCGGACGCTAGGAGCAGCATCATCACTA 899
Db 8894 TGTGATGATGTTTCAATGTTAAAGTCAATGCTGCGGACGCTAGGAGCAGCATCATCACTA 8835

OY 900 ATGTTCTCTCGAATTTAAGCAAGCAAGTGAAGAAACCAAGGAGGAGGATCAACTATGA 959
Db 8834 ACGTTTCAATTTGTTACCAACCAAGCTGTGGAAGCAACCAATGCTGAGGCTTCTGTATGA 8775

OY 960 AACCGATCAGACATATGCTCTGCTGAGTACAGGTGTCTACATTCACCTGCTACTA 1019
Db 8774 AACCAATCAGCGATATATGCTCTGAGTACAGGTGTCTACATTCACCTGCTACTA 8715

OY 1020 TCGTTACAGATGAGCCCTATATACCTCGGACAAATACCTCTGTTATTAACGTGGATA 1079
Db 8714 TGTGTAATGATATTCCTTATACATATCCGGGACAAAGCACACCTGTCTACAACTGGGATA 8655

OY 1080 GGGGCTACTTTGGCAACATACCTGCAATACGCGCTGCAACAAATCGGGAACGCCAG 1139
Db 8654 GAGCATATTTGCTATATATATCTGCAATATGCTCTTCAACAAATCAGAAATGTCACAG 8595

OY 1140 CCGTGAACACTCTAAGCAAGCTCGACACTCAACCGCGCCAAAGCTTCTTAATGCTTAG 1199
Db 8594 CCGTTGACACTTTGATATATAGTCTGCTGATATAGTATAGTATTAACCTCTTATATGCTTAG 8535

OY 1200 GAATGACTACCAAGTATTTACTACTCAATATGCCATTTCAAGTACCAACCAAGTACAG 1259
Db 8534 GTATGACTATTCAGCAAGCTGATATGCAAAAGCCATTTCAAGTATCAACCAAGTACTA 8475

OY 1260 ACAAATATGATGCAAGTGTGAAAGATGGGCTGCTAGCGCTCTTGGCAAT 1319
Db 8474 ATTAACATATGAGCAAGTGTGAAAGATGGGCTGCTAGCGCTCTTGGCAAT 8415

OY 1320 GTGCAACTTACTATTAACCAATGATATATCATTAAGTCTTATAGTATGAGAGTAAA 1379
Db 8414 GTGGCACTTACTATTAACCAATGATATATCATTAAGTCTTATAGTATGAGAGTAAA 8355

OY 1380 AAGACTTCTTAATGTGGAAGCTGCTGCATGAAGAAAGACAGACCTATATGATGACCG 1439
Db 8354 AAGACTTCTTAATGTGGAAGCTGCTGCATGAAGAAAGACAGACCTATATGATGACCG 8295

OY 1440 ACATGATGAAAAAGCTGTTGCTATGGAAGTGAAGCAAAATGCTTATGCTGGCTCC 1499
Db 8294 ACATGATGAAAAAGCTGTTGCTATGGAAGTGAAGCAAAATGCTTATGCTGGCTCC 8235

OY 1500 CTCAGGCTGTTAAACAGGAACCTCTAATCTATACAGACGAGAAATTTGAAACCAATCA 1559
Db 8234 CTCAGGCTGTTAAACAGGAACCTCTAATCTATACAGACGAGAAATTTGAAACCAATCA 8175

OY 1560 AGACCTCTCAATTTGTAGACACCTGATGACTATTTGCTGGCTATACGCGTAAATATTCAA 1619
Db 8174 AGAACACTGCTGATGTAGCTCCAGATGAATGTTGTTGTTAATACCTGTAAGTATTCCTA 8115

OY 1620 TGGCGTATGGAAGGCTATCTAACCCTGCAACCACTGTAGGCAATGGCCTTACGG 1679
Db 8114 TGGCTGATGGAAGGCTATCTAACCCTGCAACCACTGTAGGCAATGGCCTTACGG 8055

OY 1680 TCGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGGAAGAGCAATCCAGAG 1739
Db 8054 TTCGAGCAAGAAAGTTTATTCGCTCAATGATGATGATGATGATGATGATGATGATGATGATG 7995

OY 1740 ATTGGAATATACAGAGGGGCTCTACAGAAATGGAAGATTCGTTATTAATATGCTGCTC 1799
Db 7994 ACTGAGAGATGCCAGAGGAGCTTTTCAGAAACGGGGAATTTGTTATCAAAATATGAGACTC 7935

OY 1800 GTTCTAGGTGGAACCTACCTGCTCCACAAACCAACCCCATCAAGAAAGTTTCAAGCTCAT 1859
Db 7934 GCCCAATATGAGTGAACCCCTCTACTACACATCTCTCAACAGCTGAAGTTCAAGCTCAT 7875
```

OY 1860 CATGATAGTTCACTTCAAGCTCTAGCTCAACGACCTCAACCAATAATATAGTACA 1919
|
DB 7874 CATCGATAGTTCACTTCAAGCTCTAGCTCAACGACCTCAACCAATAATATAGTACA 7815
|
OY 1920 CTACCAATCTTACAAATATATAGCAACAATCAATCAATCAATCAATCAATCAATCA 1979
|
DB 7814 CTACCAATCTTACAAATATATAGCAACAATCAATCAATCAATCAATCAATCAATCA 7755
|
OY 1980 ATCTCTCAACGACACACCA 1999
|
DB 7754 ATCTCTCAACGACACACCA 7735
|

RESULT 13
US-08-116-541-3

; Sequence 3, Application US/08116541
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Pearce, Barbara J.
; APPLICANT: Toumanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/116,541
; FILING DATE: 19930901
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 960 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:
; CLONE: SPRU42
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..960
; US-08-116-541-3

Query Match 47.2%; Score 944.2; DB 5; Length 960;
Best Local Similarity 99.7%; Pred. No. 1e-256;
Matches 946; Conservatively 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TAAATCTACGACATTAATAATCACTCTGCTGACTGGGTTCTGAACGCCGCTCAA 60

|||||
DB 12 TAAATCTACGACATTAATAATCACTCTGCTGACTGGGTTCTGAACGCCGCTCAA 71
|
OY 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCTTTCTATGGAAGACA 120
|
DB 72 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCTTTCTATGGAAGACA 131
|
OY 121 TGGCTTCTTCCGACACAGGGGGATGATACCATCGGTATCCGGAGCTTTCGGCA 180
|
DB 132 TGGCTTCTTCCGACACAGGGGGATGATACCATCGGTATCCGGAGCTTTCGGCA 191
|
OY 181 TGTGCAAGCAATTTCCCTCCAGGTGATCACTCTCCACCAAGTATTAGTTATAC 240
|
DB 192 TGTGCAAGCAATTTCCCTCCAGGTGATCACTCTCCACCAAGTATTAGTTATAC 251
|
OY 241 TTTACTTTTCACTTGACCTCCGACACGATATTTCTGTAAAGCTCAGGAAGCTTGGT 300
|
DB 252 TTTACTTTTCACTTGACCTCCGACACGATATTTCTGTAAAGCTCAGGAAGCTTGGT 311
|
OY 301 AGCGATTGAGTTAGAAACAAAGCAACGACGACGAAATCTTGACCTATATATATA 360
|
DB 312 AGCGATTGAGTTAGAAACAAAGCAACGACGACGAAATCTTGACCTATATATATA 371
|
OY 361 GGTCTACATGTTCTAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 420
|
DB 372 GGTCTACATGTTCTAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 431
|
OY 421 AGACCTCAATTAATTAAGTTTACCTGAGTTAGCTTGTGCTGTGGAATGCTCAGGAC 480
|
DB 432 AGACCTCAATTAATTAAGTTTACCTGAGTTAGCTTGTGCTGTGGAATGCTCAGGAC 491
|
OY 481 AAACCAATATGACCCCTATTCACATCCAGAACGACCCGCAAGCAATTTGGTCT 540
|
DB 492 AAACCAATATGACCCCTATTCACATCCAGAACGACCCGCAAGCAATTTGGTCT 551
|
OY 541 ATCTGAAATGAATAATCAAGGCTATCTCTGCTCAACAGTGTGAAAGCAAGCTAATAC 600
|
DB 552 ATCTGAAATGAATAATCAAGGCTATCTCTGCTCAACAGTGTGAAAGCAAGCTAATAC 611
|
OY 601 ACCAATTAATGATGATCTACAAAGTCTCAATCAGCAAGTAAATTAACCTGTACATGGA 660
|
DB 612 ACCAATTAATGATGATCTACAAAGTCTCAATCAGCAAGTAAATTAACCTGTACATGGA 671
|
OY 661 TAAATTAATCAAGCAAGTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 720
|
DB 672 TAAATTAATCAAGCAAGTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 731
|
OY 721 AACTGGATGATGATCTACAAATGTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 780
|
DB 732 AACTGGATGATGATCTACAAATGTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 791
|
OY 781 TTACATACAGACGAATAGCTTGCCTATCCAGAGATGAATTCAGATGCTTACCAT 840
|
DB 792 TTACATACAGACGAATAGCTTGCCTATCCAGAGATGAATTCAGATGCTTACCAT 851
|
OY 841 TGTGATGTTTCTAAGCGTAAGTCAATTCACGCTAGAGACAGCGCATCAAGTAA 900
|
DB 852 TGTGATGTTTCTAAGCGTAAGTCAATTCACGCTAGAGACAGCGCATCAAGTAA 911
|
OY 901 TGTTCCTTCCGAAATTAACCAAGCAGTAGAACAACCAACCGCATGGGGA 949
|
DB 912 TGTTCCTTCCGAAATTAACCAAGCAGTAGAACAACCAACCGCATGGGGA 960
|

RESULT 14
PCT-US97-14436-139

; Sequence 139, Application PC/TUS9714436
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard


```

; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
; NUMBER OF SEQUENCES: 713
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/14436
; FILING DATE: 15-AUG-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,022
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50533
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
;
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; PCT-US97-14436-139
;
Query Match          40.5%; Score 808.6; DB 1; Length 2172;
Best Local Similarity 99.4%; Pred.No.5.1e-218;
Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
;
OY 1174 CGCCAAGACCTTCTTAATAGTGTAGTAATGACTACCAAGTATTCATCTACTCAATGCG 1233
    |||||||
Db 1 CGCCAAGACCTTCTTAATAGTGTAGTAATGACTACCAAGTATTCATCTACTCAATGCG 60
;
OY 1234 CATTTCAGTACACACGCAATCAGACAAAATATGAGCAAGTAGTGAAGATGGC 1293
    |||||||
Db 61 CATTTCAGTACACACGCAATCAGACAAAATATGAGCAAGTAGTGAAGATGGC 120
;
OY 1294 TGGTGGCTTAGCGTCCCTTCCAAATGCTGGAAGTCTACTATTAACCATGATATCCCTAA 1353
    |||||||
Db 121 TGGTGGCTTAGCGTCCCTTCCAAATGCTGGAAGTCTACTATTAACCATGATATCCCTAA 180
;
OY 1354 AGTCGTCTTAGTGAAGTGAAGTGAAGTCTTAATGTCGGAAGTGGTGGCATGAA 1413
    |||||||
Db 181 AGTCGTCTTAGTGAAGTGAAGTGAAGTCTTAATGTCGGAAGTGGTGGCATGAA 240
;
OY 1414 GGAAGACAGACGCTATATGATGACGACATGATGAAAAAGTCTTGAAGTGAAGTGG 1473
    |||||||
Db 241 GGAAGACAGACGCTATATGATGACGACATGATGAAAAAGTCTTGAAGTGAAGTGG 300
;
OY 1474 ACG-AAATGGCTATCTGCTTGGCTCCCTCAGGCTGTTAAACGAGACTCTAAGTATA 1592
    |||||||
Db 301 ACGAAATGGCTATCTGCTTGGCTCCCTCAGGCTGTTAAACGAGACTCTAAGTATA 360
;
OY 1533 CAGACGAGAAATGGAACACATCAAGACCTCAATTGTTGACACCTGATGAAGTAT 1592
    |||||||
Db 361 CAGACGAGAAATGGAACACATCAAGACCTCAATTGTTGACACCTGATGAAGTAT 420

```

```

OY 1593 TTGCTGCTATACGCGTAATATTCATATGCGTGTATGACAGCGCTATTTACCGCTGCA 1652
    |||||||
Db 421 TTGCTGCTATACGCGTAATATTCATATGCGTGTATGACAGCGCTATTTACCGCTGCA 480
;
OY 1653 CACACCTTAGGCAAGGCGCTTACGCTGCGTGCCTGCAAGTTTACCGCTATGATGACCT 1712
    |||||||
Db 481 CACACCTTAGGCAAGGCGCTTACGCTGCGTGCCTGCAAGTTTACCGCTATGATGACCT 540
;
OY 1713 ACCTGTGGAAGCAAGCAATCCAGAAATGGAATATATACAGAGGGGCTTACAGAAATG 1772
    |||||||
Db 541 ACCTGTGGAAGCAAGCAATCCAGAAATGGAATATATACAGAGGGGCTTACAGAAATG 600
;
OY 1773 GAGAAATGCTATTTAAATATGCTGCTGCTTCTACGTGGAATCAGTCTCCACACAAAC 1832
    |||||||
Db 601 GAGAAATGCTATTTAAATATGCTGCTGCTTCTACGTGGAATCAGTCTCCACACAAAC 660
;
OY 1833 CCCCATCACTGAAGAGTTCAGAGCTCATCATCATAGATATGTCACCTTCAAGCTACAGTCA 1892
    |||||||
Db 661 CCCCATCACTGAAGAGTTCAGAGCTCATCATCATAGATATGTCACAGTCAAGTCA 720
;
OY 1893 CCACTCCAGCACAATATATAGTACGACTACCAATCCCTAACAATATATACCAATCAAA 1952
    |||||||
Db 721 CCACTCCAGCACAATATATAGTACGACTACCAATCCCTAACAATATATACCAATCAAA 780
;
OY 1953 ATACACCCCTGATCAACAATATCAGAAATCCTCAACAGCAGACAACCA 1999
    |||||||
Db 781 ATACACCCCTGATCAACAATATCAGAAATCCTCAACAGCAGACAACCA 827

```

```

RESULT 15
US-08-911-503-139
; Sequence 139, Application US/08911503
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
; NUMBER OF SEQUENCES: 713
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,503
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,022
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50533
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
;
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 12:42:33 ; Search time 279.06 Seconds
(without alignments)
10068.144 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTGACGACATTAATAA.....ATCCTACACGACACACCA 1999

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 997330 seqs, 702755440 residues

Total number of hits satisfying chosen parameters: 1994660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_MA_New : *
1: /cgn2_6/ptodata/2/pna/PCF_NEW_COMB.seq : *
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq : *
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq : *
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq : *
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq : *
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq : *
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq : *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	4	US-08-961-083-1
2	117.2	5.9	9100	6	US-10-121-120-27
3	59.2	3.0	2370	5	US-09-540-209B-3719
4	55.2	2.8	428	5	US-09-673-476-206
5	51.8	2.6	363	5	US-09-673-476-626
6	40.6	2.0	223	5	US-09-673-476-438
7	40	2.0	610	6	US-10-027-633-5967
8	38	1.9	430	6	US-10-027-633-40693
9	37.4	1.9	242	5	US-09-789-189-1228
10	37.2	1.9	1611	7	US-60-360-038-46474
11	36.6	1.8	679	6	US-10-027-633-213202
12	36.6	1.8	2409	5	US-60-360-039-46149
13	36.4	1.8	473	5	US-09-673-476-671
14	36	1.8	689	5	US-10-027-633-148124
15	35.8	1.8	348	5	US-09-673-476-99
16	35.6	1.8	2373	5	US-09-540-209B-4136
17	35.4	1.8	8045	5	US-09-913-878A-1
18	35	1.8	510	6	US-10-027-633-82144
19	35	1.8	510	6	US-10-027-633-309095
20	35	1.8	132205	5	US-09-578-519B-25
21	34.8	1.7	3435	1	PCF-US02-10624-12
22	34.6	1.7	1247	6	US-10-027-633-254816
23	34.6	1.7	1247	6	US-10-027-633-254817
24	34.6	1.7	1247	6	US-10-027-633-254818
25	34.6	1.7	1247	6	US-10-027-633-254819
26	34.6	1.7	1247	6	US-10-027-633-254820

27	34.6	1.7	2466	5	US-09-769-744A-121	Sequence 121, App
28	34.6	1.7	2486	6	US-10-104-047-1358	Sequence 1358, Ap
29	34.2	1.7	622	6	US-10-027-633-235624	Sequence 235624,
30	34.2	1.7	747	7	US-60-360-039-33748	Sequence 32748, A
31	34.2	1.7	2463	7	US-60-360-039-26589	Sequence 26589, A
32	34.2	1.7	14803	6	US-10-105-299-11424	Sequence 11424, A
33	34	1.7	585	6	US-10-027-633-219980	Sequence 219980,
34	34	1.7	653	6	US-10-123-155-438	Sequence 438, App
35	34	1.7	653	6	US-10-137-871-438	Sequence 438, App
36	34	1.7	653	6	US-10-141-761-438	Sequence 438, App
37	34	1.7	653	6	US-10-140-864-438	Sequence 438, App
38	34	1.7	653	6	US-10-140-923-438	Sequence 438, App
39	34	1.7	653	6	US-10-141-756-438	Sequence 438, App
40	34	1.7	653	6	US-10-141-759-438	Sequence 438, App
41	34	1.7	653	6	US-10-140-472-438	Sequence 438, App
42	34	1.7	653	6	US-10-140-805-438	Sequence 438, App
43	34	1.7	653	6	US-10-142-885-438	Sequence 438, App
44	34	1.7	653	6	US-10-146-731-438	Sequence 438, App
45	34	1.7	653	6	US-10-142-426-438	Sequence 438, App

ALIGNMENTS

RESULT 1
US-08-961-083-1
Sequence 1, Application US/08961083
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1999 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-961-083-1
Query Match 100.0%, Score 1999; DB 4; Length 1999;
Best Local Similarity 100.0%, Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TAAATCTGACGACATTAATAAATCACTGCTGACTGCTGCTGACGCGCTCAA 60
|||||

Db 1 TAAATCTACGACAAATAAAAATCAATCATCTGACTTGAGGCTTCTGAAAGCCGCGTCAA 60
Oy 61 TCCCCAGCTAATGATATATCCACAGATTTGGTAAAGCAATGTTTCTATCGAAGACCA 120
Db 61 TCCCCAGCTAATGATATATCCACAGATTTGGTAAAGCAATGTTTCTATCGAAGACCA 120
Oy 121 TCGCTTCTTCGACACACAGGGGGGATGATACATCCGATCTCGGGAGCTTTCTTGCGCAA 180
Db 121 TCGCTTCTTCGACACAGGGGGGATGATACATCCGATCTCGGGAGCTTTCTTGCGCAA 180
Oy 181 TCTGAAAGCAATTTCCCTCCACAGTGGATCATCTCCACCAACAGCTTATTAAGTTGAC 240
Db 181 TCTGAAAGCAATTTCCCTCCACAGTGGATCATCTCCACCAACAGCTTATTAAGTTGAC 240
Oy 241 TTAATTTCAACTTGCAGCTTCGACACAGCTATTTCTGTAAGGCTCGAAGAGCTTGTT 300
Db 241 TTAATTTCAACTTGCAGCTTCGACACAGCTATTTCTGTAAGGCTCGAAGAGCTTGTT 300
Oy 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATATATATA 360
Db 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATATATATA 360
Oy 361 GGTCTACATGTGTATGGGAATATGGAATGAGACAGAGCTCAAACTACTATGTATA 420
Db 361 GGTCTACATGTGTATGGGAATATGGAATGAGACAGAGCTCAAACTACTATGTATA 420
Oy 421 AGACCTCATATATTAAGTTTACCTAGTTAGCTTGCTGGGTGGAAATGCTTCAGGACCC 480
Db 421 AGACCTCATATATTAAGTTTACCTAGTTAGCTTGCTGGGTGGAAATGCTTCAGGACCC 480
Oy 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGACCGCGAAATCTTGCT 540
Db 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGACCGCGAAATCTTGCT 540
Oy 541 ATCTGAATGAAATATCAAGGCTACATCTCTGTAACAGTATGAGAAAGCAAGTAAATAC 600
Db 541 ATCTGAATGAAATATCAAGGCTACATCTCTGTAACAGTATGAGAAAGCAAGTAAATAC 600
Oy 601 ACCAATTAAGTATGACATCAAAAGTCTCAAAATCAGCAAGTATTAACCTGCTTACATGA 660
Db 601 ACCAATTAAGTATGACATCAAAAGTCTCAAAATCAGCAAGTATTAACCTGCTTACATGA 660
Oy 661 TAAATTAACCTCAAGAGTATCAATCAAGTGAAGAAAGCAAGGCTTAACTACTGAC 720
Db 661 TAAATTAACCTCAAGAGTATCAATCAAGTGAAGAAAGCAAGGCTTAACTACTGAC 720
Oy 721 AACTGGATGATGATCTCTACAAATATGAGACCAAGAGCTCAAAATCAATCTGTGGGATAT 780
Db 721 AACTGGATGATGATCTCTACAAATATGAGACCAAGAGCTCAAAATCAATCTGTGGGATAT 780
Oy 781 TTACAAATACAGCAATACGTTGCTATCCAGAGCATGAATGCAAGTCGCTTACCAT 840
Db 781 TTACAAATACAGCAATACGTTGCTATCCAGAGCATGAATGCAAGTCGCTTACCAT 840
Oy 841 TGTGATGATTTCTTAAGGTAAGTATGCGCAGCTAGAGAGCCCATCAATCAAGTAA 900
Db 841 TGTGATGATTTCTTAAGGTAAGTATGCGCAGCTAGAGAGCCCATCAATCAAGTAA 900
Oy 901 TGTTCCTTCGGAATTAACCAAGAGTGAAGAAACAAACCGCGATCGGATCAACTATATA 960
Db 901 TGTTCCTTCGGAATTAACCAAGAGTGAAGAAACAAACCGCGATCGGATCAACTATATA 960
Oy 961 ACCGATACAGACTATGCTGCTGCTGGAGTAGGGTCTATCAATCAAGTCACTACTAT 1020
Db 961 ACCGATACAGACTATGCTGCTGCTGGAGTAGGGTCTATCAATCAAGTCACTACTAT 1020
Oy 1021 CGTTCAGATGAGCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
Db 1021 CGTTCAGATGAGCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
Oy 1081 GGGCTACTTTGGCAACATACACCTTGCATATAGCCCTGCAACAAATGCGAAACGCTCCAGC 1140
Db 1081 GGGCTACTTTGGCAACATACACCTTGCATATAGCCCTGCAACAAATGCGAAACGCTCCAGC 1140

Oy 1141 CGTGAACCTTAACAAAGTGGAGTCAACCGCGCAAGACTTTCTAAATGCTTAGG 1200
Db 1141 CGTGAACCTTAACAAAGTGGAGTCAACCGCGCAAGACTTTCTAAATGCTTAGG 1200
Oy 1201 AATGCACTACCCAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
Db 1201 AATGCACTACCCAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
Oy 1261 CAAAAATATGAGCAAGTATGAGAAAGAGTGGCTGCTTACGCTTGGCAATAGG 1320
Db 1261 CAAAAATATGAGCAAGTATGAGAAAGAGTGGCTGCTTACGCTTGGCAATAGG 1320
Oy 1321 TGGAACTTACTTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
Db 1321 TGGAACTTACTTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
Oy 1381 AGAGTTCTTAATGAGCAACTGTCCTATGAGAAAGAGAGAGAGCTTATATGATGACCGA 1440
Db 1381 AGAGTTCTTAATGAGCAACTGTCCTATGAGAAAGAGAGAGAGCTTATATGATGACCGA 1440
Oy 1441 CATGATGAAACAGTCTGATTAATGAGAACTGAGAAATGCGTATGCTTGGCTGCC 1500
Db 1441 CATGATGAAACAGTCTGATTAATGAGAACTGAGAAATGCGTATGCTTGGCTGCC 1500
Oy 1501 TCAGGCTGGTAAACAGAACTTAACTATACAGAGAGAAATTTGAACCAACATCAA 1560
Db 1501 TCAGGCTGGTAAACAGAACTTAACTATACAGAGAGAAATTTGAACCAACATCAA 1560
Oy 1561 GACCTCTCAATTTGATGACACTGATGAACATATTTGCTGCTATACGCTTAATATTCAT 1620
Db 1561 GACCTCTCAATTTGATGACACTGATGAACATATTTGCTGCTATACGCTTAATATTCAT 1620
Oy 1621 GCGCTATGAGCAAGGCTATTCACACGCTGACACACTGTAAGCAATGCGCTTAGCGT 1680
Db 1621 GCGCTATGAGCAAGGCTATTCACACGCTGTAACGCTGTAAGCAATGCGCTTAGCGT 1680
Oy 1681 CGCTGCCAAAGTTTACGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1681 CGCTGCCAAAGTTTACGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Oy 1741 TTGGAATATACAGAGGCGCTCTACAGAAATGAGAAATTCGATTTAAATATGCTGCTG 1800
Db 1741 TTGGAATATACAGAGGCGCTCTACAGAAATGAGAAATTCGATTTAAATATGCTGCTG 1800
Oy 1801 TTCTAGTGGAACTACCTGCTCCACAACACCCCAATCACTGAAGTCAAGCTCATC 1860
Db 1801 TTCTAGTGGAACTACCTGCTCCACAACACCCCAATCACTGAAGTCAAGCTCATC 1860
Oy 1861 ATCAGATAGTCACTTACAGAGTCAAGCTCAACCAAGCAACAAATTAATAGTAGAC 1920
Db 1861 ATCAGATAGTCACTTACAGAGTCAAGCTCAACCAAGCAACAAATTAATAGTAGAC 1920
Oy 1921 TACCAATCTTAACAAATTAATGCAACAAATTAATGCAACCAATTAATGCAACAAATTAATGCA 1980
Db 1921 TACCAATCTTAACAAATTAATGCAACAAATTAATGCAACCAATTAATGCAACAAATTAATGCA 1980
Oy 1981 TCCTCAACGACACACCA 1999
Db 1981 TCCTCAACGACACACCA 1999

RESULT 2
US-10-121-120-27/c
; Sequence 27, Application US/10121120
; GENERAL INFORMATION:
; APPLICANT: Bergeron, Michel G.
; APPLICANT: Ouellette, Marc
; APPLICANT: Roy, Paul H.
; TITLE OF INVENTION: Specific and Universal Probes and Amplification
; TITLE OF INVENTION: Primers
; TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial
; TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical

;; TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro
;; FILE REFERENCE: 12287.31
;; CURRENT APPLICATION NUMBER: US/10/121,120
;; CURRENT FILING DATE: 2002-04-11
;; PRIOR APPLICATION NUMBER: 09/452,599
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: 08/304,732
;; PRIOR FILING DATE: 1994-09-12
;; NUMBER OF SEQ ID NOS: 177
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO: 27
;; LENGTH: 9100
;; TYPE: DNA
;; ORGANISM: Haemophilus influenzae
US-10-121-120-27

Query Match 5.9%; Score 117.2; DB 6; Length 9100;
Best Local Similarity 50.3%; Pred. No. 1.5e-22;
Matches 349; Conservative 0; Mismatches 333; Indels 12; Gaps 2;

QY 22 TCACTCATGCTGACTGGGTTCTGACGCGCGTCAATGCCCAAGCTAATGATATCC 81
DB 3498 TAAATTAATGGCGAAGTGGGAGACGCGGTATTCAGTGAATAGCCGATGCGC 3439
QY 82 CACAGATTGGTTAAGGCAATGTTCTATCGAAGACATCGCTTCTTGACACAGGG 141
DB 3438 ACAACGCTTAATGAGCGATTTTACGAGGAAACAGTCTGTTTACGATCATCG 3379
QY 142 GATTCATACATCCGATCTGAGGAGCTTCTGCGCAATCTGCAAA---GCAATTCCT 198
DB 3378 ATTAGCCCTATCGGATTCGCCGTCATGTTTGTGCGAGATGATAGGCGGTGATC 3319
QY 199 CCAAGGTGATCACTCTACCCACAGTTGATTAAGTACTTACTTTCAACTTCGAC 258-
DB 3318 ACAAGCGCGAATGATGATTAACAATTAAGGCGTAACTTTCTTA-----AC 3268
QY 259 TTCCGACGAGTATTTCTGTAAGGCTCAGAAAGCTTGTTAGCATTCAGTTAGACA 318
DB 3267 CTCAGAAAAACCATTAATTCGTAAAGCTGTGAAGCGGTGGTGGTAATCGAAAA 3208
QY 319 AAAAGCAACCAAGCAAGAAATCTTGACTATATATAAAGTCTACATGCTAATGG 378
DB 3207 TACGTCAACAAACAAAGAAATATAGCTTTATTTAAACAAATCTTTTGGCTATG 3148
QY 379 GAATATGGAATGACAGACAGCTCAAACTAATGTTAAAGACCTCAATTAATTAG 438
DB 3147 TTCTATGTTGTGACGCGGACACAACTATTCGTAATATCAATGATGATTTGAC 3088
QY 439 TTACTCTAGTACCTTCTGCTGCTGCTGGAATGCTCAGGACCAACCAATATGACCC 498
DB 3087 CTTATGGAAATGGGATTAATGCTGTTTACCTAAAGCACTTCAACATGAACCCGCT 3028
QY 499 TTACATCCAGAAAGAGCCAGACCGCGAAACTTGCTTATCTGAATGAAATAATCA 558
DB 3027 TTAATCTTTAAAGCTTGAAGAGACCGCGCAATGTGGTCTAAGCCGATGTATATGA 2968
QY 559 AGGCTACATCTCTGCTGAACAGTATGAAAGACAGTCAATACCAATTACTGATGACT 618
DB 2967 AAAATACATCAGCAAGAAAGAAATATGATGCTGATGAAAGAGCCGATTTGGCGAGCTA 2908
QY 619 ACAAGTCAATCAATGACGAAGTATACCTGCTTACATGATTAATACCTCAGGAAGT 678
DB 2907 TCACGGCGGCAAAATTTGAATTTTCAGCCGATTAATGCTCACTGAATGTCGTCACAGAAAT 2848
QY 679 CATCATCAAGTTGAAGAGAAACAGGCTATTAAC 712
DB 2847 GGTGCTGCTTTTGGCGAAGAAATGCTTACACC 2814

RESULT 3
US-09-540-209B-3719
; Sequence 3719, Application US/09540209B

;; GENERAL INFORMATION:
;; APPLICANT: Gary L. Breton
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
;; FILE REFERENCE: 2709.1001-001
;; CURRENT APPLICATION NUMBER: US/09/540,209B
;; CURRENT FILING DATE: 2000-04-04
;; NUMBER OF SEQ ID NOS: 10444
;; SEQ ID NO: 3719
;; LENGTH: 2370
;; TYPE: DNA
;; ORGANISM: B. fragilis
US-09-540-209B-3719

Query Match 3.0%; Score 59.2; DB 5; Length 2370;
Best Local Similarity 46.9%; Pred. No. 2.2e-06;
Matches 261; Conservative 0; Mismatches 283; Indels 12; Gaps 2;

QY 42 GTTCTGACCCCGCTCAATGCCCAAGCTAATGATATTCGCCAGATTGGTTAGGCA 101
DB 233 gtaagaaatcgtgatacaccctataataatgaacttaccacaacatgtaacatgcat 292
QY 102 TCGTTCTATCGAAGACATCGCTTCTGACACAGGCGGATTTGATACCATCGTATCC 161
DB 293 tgatgcacagagagacgttcgttaccagaaatccggtatcgatgcgaagcgctga 352
QY 162 TGGGAGCTTCTT-----GCGCAATCTGCAAAAGCAATTCCTCCAGGTGATCAATC 215
DB 353 taatgctatgtaagagcgtgatactgatacgagaataatgagtgagtgagagagacac 412
QY 216 TCACCAACAGTGTATTA-----GTGACTTACTTTTCACTTGACCTTCGACACAGA 269
DB 413 ttccacaagaagcgtccgaagcaatglttacgagcgaagltgccaagaataacgctgcagc 472
QY 270 CTAATTCCTGTAAGGCTCAGAAAGCTTGTGATGATGATGATGATGATGATGATGATG 329
DB 473 gctgtttcagaagacgataagatgagtgatgctgacgtaaaactggaactatatacaa 532
QY 330 AGCAGAAATCTTGACCTATATATAAAGTCTACATGTTGGAATGGAACTATGAA 389
DB 533 aggaagaaattttgagatgatacctaataatttgacttcgtaataatgacgtagaa 592
QY 390 TGCAGACAGAGCTCAAACTATGTAAGACCTCAATTAATTAAGTTTACCTCAGT 445
DB 593 ttaaaacgcttataatactatcttcgatacgaaaccaaagatcgaataatagaacag 652
QY 450 TAGCTTGTGCTGATGATGCTTCCAGGACCAACCAATATGACCCCTATTCATCAG 509
DB 653 cgtctacgtatgatacgtatgataaataccttcgcttacaatccggtgcgtcaag 712
QY 510 AAGCAGCCCAAGCCCGCAAACTTGCTTATCTGAATGAAATAATCAAGGCTACATCT 569
DB 713 aggttcgcgcggaagcgaataacagtgatagataaatgagaagccggtatataca 772
QY 570 CTGCTGAACAGTATGA 585
DB 773 cagcgagaagagtgta 788

RESULT 4
US-09-673-476-206
; Sequence 206, Application US/09673476
;; GENERAL INFORMATION:
;; APPLICANT: COLE, STEWART
;; APPLICANT: BUCHRISEB-BROSCH, ROLAND
;; APPLICANT: GORDON, STEPHEN
;; APPLICANT: BILAUDT, ALAIN
;; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
;; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
;; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
;; FILE REFERENCE: 05394.0011-00000

```

: CURRENT APPLICATION NUMBER: US/09/673.476
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: PCT/IB99/00740
: PRIOR FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: 09/060,756
: PRIOR FILING DATE: 1998-04-16
: NUMBER OF SEQ ID NOS: 743
: SOFTWARE: PatentIn Ver. 2.2
: SEQ ID NO 206
: LENGTH: 428
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (280)
: OTHER INFORMATION: a, t, c or g
US-09-673-476-206
```

```

Query Match          2.8%; Score 55.2; DB 5; Length 428;
Best Local Similarity 50.8%; Pred. No. 1.5e-05;
Matches 132; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
```

```

OY 281 AAGGCTCAGGAAGCTTGGTTAGCGATTGATAGCAAAAAGCAAGCAAGCAAAATC 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 aagctcggagatccgagatgcactcagcgcgcaagaagcctcacaatactgaatc 66

OY 341 TTGACCTACTATATTAATAGGTCTACATGCTATGGAAGTAATGAGTACAGACGA 400
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 ctgaccgcatctgtaacctgctcgtcgcaataactcgttcgcaagaagcgcg 126

OY 401 GCTCAAACTACTATGTAAGAGCTCAATATTAAGTTTACCTACGTTAGCTTGGCTG 460
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 ggcgaagaagctacttgcgcatacaagcgctcgaacctgaattgcaagaagcgctgctg 186

OY 461 GCTGAATGCTCTAGGACGACCAACCAATATGACCCCTATTCATTCAGAACGACCCCA 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 gccgcatactgtaactgcacgcagcgcctcaaccgctacacacccgcgagcgcgctg 246

OY 521 GACCCGCCAACTTGTGCTT 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 gccgcgagcgaacgtgctct 266
```

```

RESULT 5
US-09-673-476-626
: Sequence 626, Application US/09673476
: GENERAL INFORMATION:
: APPLICANT: COLE, STEWART
: APPLICANT: BUCHRIEISER-BROSCH, ROLAND
: APPLICANT: GORDON, STEPHEN
: APPLICANT: BILLAUFT, ALAIN
: TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
: TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
: TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
: FILE REFERENCE: 05394.0011-00000
: CURRENT APPLICATION NUMBER: US/09/673.476
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: PCT/IB99/00740
: PRIOR FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: 09/060,756
: PRIOR FILING DATE: 1998-04-16
: NUMBER OF SEQ ID NOS: 743
: SOFTWARE: PatentIn Ver. 2.2
: SEQ ID NO 626
: LENGTH: 363
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
US-09-673-476-626
```

```
Query Match          2.6%; Score 51.8; DB 5; Length 363;
```

```

Best Local Similarity 52.6%; Pred. No. 0.00013;
Matches 113; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

OY 326 ACCAAGCAAGAAATCTGACCTACTATATTAATAGTCTACATGCTAATGGAAGATAT 385
    ||| ||| ||||| ||||| ||| ||| ||||| ||| ||| ||| ||| ||| |||
Db 26 acaaatctgaaatcttccatccgatacttgaacctgtctcgttgcgaataactcgttc 85

OY 386 GGAATGCAGACAGACGCTCAAAATCTACTATGTTAAAGACCTCAATATTAATTAAGTTTACT 445
    ||| ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 ggcgcgcgaagcgcgcgcgaacacgtaacttcggaatcaacgcgtccgaactgaatggcag 145

OY 446 CAGTTAGCCTTGCTGGTGAATGCTTCAGGACGACCAACCAATATGACCCCTATTTCACAT 505
    ||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 caagcgcgctcgtgcgcgcatggtgcaatcgaccagcagcctcaaccgctacacccaac 205

OY 506 CCAGAAGCAGCCCAAGACCCGCGCAACTTGTGCTT 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 cccgacgcgcgcgtgcgcgcgcgcgcgcgcgtgctct 240
```

```

RESULT 6
US-09-673-476-438
: Sequence 438, Application US/09673476
: GENERAL INFORMATION:
: APPLICANT: COLE, STEWART
: APPLICANT: BUCHRIEISER-BROSCH, ROLAND
: APPLICANT: GORDON, STEPHEN
: APPLICANT: BILLAUFT, ALAIN
: TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
: TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
: TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
: FILE REFERENCE: 05394.0011-00000
: CURRENT APPLICATION NUMBER: US/09/673.476
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: PCT/IB99/00740
: PRIOR FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: 09/060,756
: PRIOR FILING DATE: 1998-04-16
: NUMBER OF SEQ ID NOS: 743
: SOFTWARE: PatentIn Ver. 2.2
: SEQ ID NO 438
: LENGTH: 223
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (136)
: OTHER INFORMATION: a, t, c or g
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (158)
: OTHER INFORMATION: a, t, c or g
: NAME/KEY: modified_base
: LOCATION: (203)
: OTHER INFORMATION: a, t, c or g
US-09-673-476-438
```

```

Query Match          2.0%; Score 40.6; DB 5; Length 223;
Best Local Similarity 48.6%; Pred. No. 0.16;
Matches 106; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

OY 281 AAGGCTCAGGAAGCTTGGTTAGCGATTGATAGCAAAAAGCAAGCAAGCAAAATC 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 aagctcggagatccgagatgcactcagcgcgcaagaagcctcacaatactgaatc 65

OY 341 TTGACCTACTATATTAATAGGTCTACATGCTATGGAAGTAATGAGTACAGACGA 400
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 ctgaccgatactgtaacctgctcgtcgcaataactcgttcgcaagaagcgcg 125

OY 401 GCTCAAACTACTATGTAAGAGCTCAATATTAAGTTTACCTACGTTAGCTTGGCTG 460
```


PRIOR APPLICATION NUMBER: US 60/161,363

125 TCTTCGACCAAGGGGATGGATACCAATCCGTA

[illegible]

Db 201 TTGCGGACACAGCGGGGTGACTGGAAGGGCAACCTG 163

Search completed: June 13, 2002, 14:47:52
Job time: 7519 sec

• • •

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 03:12:54 ; Search time 61.45 Seconds
(without alignments)
1203.827 Million cell updates/sec

Title: US-08-961-083-2
Perfect score: 3484
Sequence: 1 KIVDNKNQLIADLGSRRVN.....TQSSNTTPOQNONPQAPQ 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*

1:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	666	19	AAW55063
2	3466	99.5	682	17	AAW04359
3	3466	99.5	719	22	AAU37830
4	2208	63.4	420	20	AAV56106
5	1751	50.3	778	22	AAU35135
6	1621	46.5	320	16	AAK70153
7	1361	39.1	462	22	AAU03646
8	817.5	23.5	727	22	AAU34286
9	817.5	23.5	727	22	AAU37184
10	640	18.4	805	21	AAV5603
11	638	18.3	805	21	AAV5602

12	636	18.3	731	19	AAW44849	S. pneumoniae peni
13	635.5	18.2	805	21	AAV75601	Neisseria gonorrhe
14	625	17.9	660	18	AAW55545	H. pylori ORF 06ep
15	625	17.9	660	18	AAW55576	H. pylori ORF 06ep
16	608.5	17.5	596	18	AAW89836	Protein encoded by
17	586	16.8	812	22	AAW78604	Lawsonia intracell
18	567	16.3	828	17	AAW04357	E. coli penicillin
19	555	15.9	850	17	AAW86955	E. coli PBP 1A tra
20	552	15.8	774	22	AAU36453	Pseudomonas aerugi
21	538.5	15.5	536	20	AAW89885	Antigen 2 from cin
22	529.5	15.2	821	21	AAV81757	Streptococcus pneu
23	526	15.1	823	17	AAW04358	E. coli penicillin
24	526	15.1	844	22	AAU34446	E. coli cellular p
25	525.5	15.1	781	22	AAU36684	Hemophilus influe
26	519.5	14.9	846	22	AAU38311	Salmonella typhi c
27	514	14.8	844	17	AAW86952	E. coli PBP 1B tra
28	512	14.7	844	17	AAW86953	E. coli PBP 1B tra
29	508	14.6	853	22	AAU55400	Propionibacterium
30	489	14.0	836	17	AAW86954	E. coli PBP 1B tra
31	457	13.1	447	18	AAW20510	H. pylori inner me
32	457	13.1	799	22	AAU49462	Propionibacterium
33	450.5	12.9	770	22	AAW98353	Escherichia coli p
34	441.5	12.7	795	22	AAW90056	C. glutamicum prote
35	437.5	12.6	678	22	AAW81107	Mycobacterium tube
36	422	12.1	720	22	AAW93052	C. glutamicum prote
37	419	12.0	810	22	AAW81232	Mycobacterium tube
38	377	10.8	344	18	AAW55362	H. pylori ORF 14gp
39	377	10.8	344	18	AAW20917	H. pylori cell env
40	370.5	10.6	532	17	AAW86957	E. coli penicillin
41	370.5	10.6	553	17	AAW86956	E. coli penicillin
42	341.5	9.8	231	20	AAW90230	R. eutropha Mgt pa
43	288.5	8.3	269	19	AAW61395	Staphylococcus aur
44	287	8.2	338	22	ABG24982	Novel human diagno
45	280.5	8.1	269	22	AAW81893	S. epidermidis ope

ALIGNMENTS

RESULT 1
AAW55063
AAW55063 standard; Protein; 666 AA.

ID	AAW55063;
XX	
AC	02-OCT-1998 (first entry)
XX	
DT	
XX	
DE	Streptococcus pneumoniae SP001 protein.
XX	
KW	Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
KW	detection; pneumonia; otitis media; meningitis.
XX	
OS	Streptococcus pneumoniae.
XX	
PN	W09818930-A2.
XX	
PD	07-MAY-1998.
XX	
PF	30-OCT-1997; 97WO-US19422.
XX	
PR	31-OCT-1996; 96US-0029960.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Choi GH, Hromockyj A, Johnson LS, Kunsch CA;
XX	
DR	WPI: 1998-272224/24.
XX	
DR	N-PSDB; AAV27523.
XX	
PT	Nucleic acid encoding antigenic peptide(s) from Streptococcus
XX	pneumoniae - or their epitope-containing fragments, useful in
XX	protective or therapeutic vaccines, and for diagnosis

PS Claim 11: Page 48; 118pp; English.

XX The present sequence represents a protein from *Streptococcus pneumoniae*.
 CC The nucleic acid sequence encoding the *Streptococcus pneumoniae* protein
 CC can be useful in vaccines for inducing protective antibodies against
 CC *Streptococcus pneumoniae*, for treatment or prevention of infection e.g.
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
 CC are used to detect *Streptococcus* infection (by usual hybridisation or
 CC amplification methods), also for isolating *Streptococcus* genes or their
 CC allelic variants. The protein can be used similarly to detect specific
 CC antibodies in standard immunoassays, especially for diagnosing or
 CC monitoring infections. Antibodies which bind the protein are used to
 CC detect corresponding antigens, to purify the protein and for passive
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000
 CC (especially 10-300) mu g/ml per dose.

XX Sequence 666 AA;

Query Match Best Local Similarity 100.0%; Score 3484; DB 19; Length 666;

Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KIYNKNOIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAPLRN 60
    |||
DB 1 KIYNKNOIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAPLRN 60
    |||

QY 61 LOSNSLOGSFTLTOOLIKLTFSTSTSDQTSRKQAEWALIQLEOKATKQEIILTYINK 120
    |||
DB 61 LOSNSLOGSFTLTOOLIKLTFSTSTSDQTSRKQAEWALIQLEOKATKQEIILTYINK 120
    |||

QY 121 VYMSGNGYGMOTAAONYGKDLNNLSLPOLALAGMPAPNOQYDPYSHPEAAODRNLVL 180
    |||
DB 121 VYMSGNGYGMOTAAONYGKDLNNLSLPOLALAGMPAPNOQYDPYSHPEAAODRNLVL 180
    |||

QY 181 SEMKNGYISAQYERKAVNTPITDGLQSLKASANYPAYMDNYLKEVINQVEERTGYNLLT 240
    |||
DB 181 SEMKNGYISAQYERKAVNTPITDGLQSLKASANYPAYMDNYLKEVINQVEERTGYNLLT 240
    |||

QY 241 TGMQVYTNVDDAOKHMDIYNTDEYVAYPPDELOVASTIYDVSNGKIYIOLGARNHGSN 300
    |||
DB 241 TGMQVYTNVDDAOKHMDIYNTDEYVAYPPDELOVASTIYDVSNGKIYIOLGARNHGSN 300
    |||

QY 301 VSFQINQAVENRPMGSTMKPTITDYPALAEYGVDSATPIVHDEPYNPGTTPPYNMMDR 360
    |||
DB 301 VSFQINQAVENRPMGSTMKPTITDYPALAEYGVDSATPIVHDEPYNPGTTPPYNMMDR 360
    |||

QY 361 GYFGNITLQVALQOQRNYPAVETLNKVGILNRAKTFPLNGIDIDPSIHSNMAISSNTTESD 420
    |||
DB 361 GYFGNITLQVALQOQRNYPAVETLNKVGILNRAKTFPLNGIDIDPSIHSNMAISSNTTESD 420
    |||

QY 421 KKYGASSEKMAAAYAFANGSTYRKPMIIRKVVVSDGSEKESNNGTRAMKETTAYMMTD 480
    |||
DB 421 KKYGASSEKMAAAYAFANGSTYRKPMIIRKVVVSDGSEKESNNGTRAMKETTAYMMTD 480
    |||

QY 481 MMKTYLITGTRNMYLAWLPQAGKGTSTNYDEETENHIKTSQFAPELFCAGYRKRTSM 540
    |||
DB 481 MMKTYLITGTRNMYLAWLPQAGKGTSTNYDEETENHIKTSQFAPELFCAGYRKRTSM 540
    |||

QY 541 AVMTGYSNRLPLVGNGLTVAAYKVRSMKTYLSEGSNPEDNINIPGLRNGEFVEKNAR 600
    |||
DB 541 AVMTGYSNRLPLVGNGLTVAAYKVRSMKTYLSEGSNPEDNINIPGLRNGEFVEKNAR 600
    |||

QY 601 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOOSNTTPDOQON 660
    |||
DB 601 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOOSNTTPDOQON 660
    |||

QY 661 POPAOP 666
    |||
DB 661 POPAOP 666
    |||

```

RESULT 2

ID AAM04359 standard; Protein: 682 AA.

AC AAM04359;

DT 03-DEC-1996 (first entry)

DE S. pneumoniae penicillin binding protein 1A soluble variant.

KW Penicillin binding protein; PBP 1A; bifunctional protein;

KW transglycosylase; transpeptidase; identification; assay; inhibitor;

KW antibiotic resistant; bacteria; soluble variant; protein structure;

KW X-ray crystallography; determination.

OS Streptococcus pneumoniae.

PN GB2290792-A.

PD 10-JAN-1996.

PE 29-JUN-1995; 95GB-0013306.

PR 24-NOV-1994; 94SE-0004072.

PR 01-JUL-1994; 94IN-0000580.

PA (ASTR) ASTRA AB.

PI Balganesch TS, Town CM;

DR WPI; 1996-042232/05.

DR N-PSDB; AAT08027.

PS Sol. derivs. of bifunctional penicillin binding protein (BPP) -

PT opt. lack transglycosylase activity; useful to identify and assay

PT for antibodies or cpds. which bind BBPs

PS Claim 2; Pages 65-68; 108pp; English.

CC The present sequence is a soluble variant of the S. pneumoniae

CC penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino

CC acid residues of the wild type protein. Wild type PBP is a

CC bifunctional protein, which binds the cell membrane when expressed

CC in a bacterial cell, having transglycosylase and transpeptidase

CC activities. The variant protein (NCIMB 40665) in conjunction with

CC a labelled anti-bifunctional PBP monoclonal antibody, can be used

CC to identify and assay for cpds. which bind bifunctional PBP. Such

CC cpds., as inhibitors of bifunctional PBP have a potential use in

CC therapeutic cpds. which inhibit the growth of antibiotic resistant

CC bacteria. The soluble variant may also be used in X-ray

CC crystallography.

XX Sequence 682 AA;

Query Match Best Local Similarity 99.5%; Score 3466; DB 17; Length 682;

Matches 662; Conservative 99.4%; Pred. No. 2e-233; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 KIYNKNOIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAPLRN 60
    |||
DB 17 KIYNKNOIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAPLRN 76
    |||

QY 61 LOSNSLOGSFTLTOOLIKLTFSTSTSDQTSRKQAEWALIQLEOKATKQEIILTYINK 120
    |||
DB 61 LOSNSLOGSFTLTOOLIKLTFSTSTSDQTSRKQAEWALIQLEOKATKQEIILTYINK 136
    |||

QY 121 VYMSGNGYGMOTAAONYGKDLNNLSLPOLALAGMPAPNOQYDPYSHPEAAODRNLVL 180
    |||
DB 121 VYMSGNGYGMOTAAONYGKDLNNLSLPOLALAGMPAPNOQYDPYSHPEAAODRNLVL 196
    |||

QY 181 SEMKNGYISAQYERKAVNTPITDGLQSLKASANYPAYMDNYLKEVINQVEERTGYNLLT 240
    |||

```

Db	197	semkrqgylaseqjekavntpildqglksaasnypaymdyilkevlnqyeeegynllt	256
Qy	241	TGMDVYTNWDQDAOKHLMIDYINTDEYVAYPPDDELQVASTTIVDSNGKYIAQLGARHGSN	300
Db	257	tgmdivtnvdqeaqqlwdiynltdeyaypddelqvasltivdangkvilqarhgsn	316
Qy	301	VSFQINQAVETNRWGSMTKRIPTDYAPALEGYVDSTATIVHDEPYNYPGTNPVYMWDR	360
Db	317	vsfqlngavevnrwgsmtkrlptdyapaleyvyvestaltlvhdepynpynlnpvywnwr	376
Qy	361	GYFGNITTLQVALQOOSRNPVAVETLNKVGILNAKTFELNGLGIDYPSIHYSNAISNTESD	420
Db	377	gyfgnltlqvalqgsrnpvavetlnkvglmkactflnglqldydpshysnaissntesd	436
Qy	421	KRYGASSEKMAAAYAAFAANGSTYKPMYIHKVPSDSEKEFSNVGTRAKETVAYMMTD	480
Db	437	krygassekmaayaafangstyykpmylhkvpsdsgekefsnvgtrankettaymmtd	496
Qy	461	MMKTYLVTGTGRNAVLAFLPDAGTGTGNSNTDEIEHNIKTSQVAPDELPACTTKRYSM	540
Db	497	mmktylvgtgrnavlamlfpdactgtgnsntdeiehniktsqvapdelpacttkrysm	556
Qy	541	AAVNGYSRRLPLVLNGGTLTVAAKVYRSMMFTYLSGSPEDMNIPEGLYRNGEYFKNGAR	600
Db	557	avvngysrrlplvlnngtlvtaakvyrsmtfytlsgsnpedmnipeglyrngeyfkngar	616
Qy	601	STWNSPAPQDPSTRESSSSSDSSSTSOSSSTTPSTNNSTTTPNNNNTOQSNTTPDOQON	660
Db	617	stwsnpapqpdpstresssssdssstsoasssttpstnnsttppnnnntoqsnttpdoqon	676
Qy	661	PQPAQP 666	
Db	677	pqpapq 682	
RESULT 3			
AAU37830			
ID	AAU37830	standard; Protein: 719 AA.	
XX	AAU37830;		
XX	14-FEB-2002 (first entry)		
De	Streptococcus pneumoniae cellular proliferation protein #259.		
XX	Antisense: prokaryotic cellular proliferation protein;		
KW	antibiotic; antibacterial; drug design.		
XX	Streptococcus pneumoniae.		
XX	MO200170955-A2.		
PD	27-SEP-2001.		
XX	21-MAR-2001; 2001MO-USO9180.		
XX	21-MAR-2000; 2000US-191078P.		
PR	23-MAY-2000; 2000US-206848P.		
PR	26-MAY-2000; 2000US-207727P.		
PR	23-OCT-2000; 2000US-242578P.		
PR	27-NOV-2000; 2000US-253625P.		
PR	22-DEC-2000; 2000US-257931P.		
PR	16-FEB-2001; 2001US-269308P.		
XX	(ELIT-) ELITRA PHARM INC.		
XX	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;		
PI	Yamamoto RT, Xu HH,		
XX	WPI: 2001-611495/70.		
DR	N-PSDB: AAS55689.		
XX	New polynucleotides for the identification and development of		

PT		antibiotics, comprise sequences of antisense nucleic acids -
XX		
PS		
XX		
XX		
CC	Example 3; Seq ID No 13423; 51pp; English.	
CC	The invention relates to antisense inhibitors of genes essential to	
CC	prokaryotic cellular proliferation, their use in identifying the	
CC	genes, their use in the discovery of novel antibiotics, the essential	
CC	genes themselves and the encoded proteins. The prokaryotes used are	
CC	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella	
CC	pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The	
CC	invention is also useful for the identification of potential new targets	
CC	for antibiotic development. The antisense nucleic acids can also be used	
CC	to identify proteins used in proliferation, to express these proteins,	
CC	and to obtain antibodies capable of binding to the expressed proteins.	
CC	The proteins can be used to screen compounds in rational drug discovery	
CC	programmes. The antisense nucleic acid sequence is also useful to screen	
CC	for homologous nucleic acids which are required for cell proliferation in	
CC	a wide variety of organisms. The present sequence represents an	
CC	essential prokaryotic cellular proliferation protein.	
CC	Note: The sequence data for this patent did not form part	
CC	of the printed specification, but was obtained in electronic	
CC	format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences.	
CC		
XX		
SQ	Sequence 719 AA:	
	Query Match 99.5%; Score 3466; DB 223; Length 719;	
	Best Local Similarity 99.4%; Pred. No. 2.le-23;	
	Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
OY	1 KIYDNKNOLADIAGSERRYNAQAINDPDLVKAIVSIEDHREDFRGIDITRIILGAPLRN 60	
DB	54 Kiydnknqniadiagserrynagaandi p d l v k a i v s i e d h r f d r g i d i t r i i l g a f l r n 113	
OY	61 LOSNSLOGGSTLTQOLIKLTFSTSTSDQTISRKQEAMLAIQLQKATKDEILTYINK 120	
DB	114 l o s n s l q g s s a l t q d l k l t f s t s d q t i s r k a e a m l a i q l e q a t k g e i l l y u n k 173	
OY	121 VMSNGNNGMOTAAONNYGXKDNLNLSLPOLALIAAMPAPNPQNDYRSHPEAODRRNLVL 180	
DB	174 v y m s n g n g m o t a a o n n y g x k d n l n l s p o l a l i a g m p a n p q n d y s h p e a o d r r n l v l 233	
OY	181 SEMKNQGYISADOUYRKAVWTPRTDGLQSILKSASNPAYMDNKLAKVINQVEETGYNLLT 240	
DB	234 s e m k n g y i s a d o u y r k a v w t p r t d g l q s i l k s a s n p a y m d n k l a k v i n q v e e t g y n l l t 293	
OY	241 TGMADVTVNVDOEAKNLMWDIYNFDEVAYRPDELQOVASTIVDVNSGVYIAQLGARHGSN 300	
DB	294 t g m d v t y n v d o e a k n l m w d i y n f d e v a y r p d e l q o v a s t i v d v n s g v y i a q l g a r h g s n 353	
OY	301 VSFGINQAVENTRDMGSTMKPRTTDVARALEGVNDSTATIVNDEXYNPGCTPYRYNMNR 360	
DB	354 v s f g i n q a v e n t r d m g s t m k p r t t d v a r a l e g v n d s t a t i v n d e x y n p g c t p y r y n m n r 413	
OY	361 GFYGNTITLOVALQOQRNNPAYVELLNKKVGNLRKKTPLNGLGIDYPSIIHSNATSMTESD 420	
DB	414 g y f g n t i t l o v a l q o q r n n p a y v e l l n k k v g n l r k k t p l n g l g i d y p s i i h s n a t s m t e s d 473	
OY	421 KKYGASSEKMAAVALAFANGCTYYKPMYIHKVFSDSGESEKSVNSTRAMKETTAVMMD 480	
DB	474 k k y g a s s e k m a a v a l a f a n g c t y y k p m y i h k v f s d s g e s e k s v n s t r a m k e t t a y m m d 533	
OY	481 MKKYTLVTGTGNANLAFLMPQAGTKGTGSNYNDELENIHTSQFPAPBEPFAGYRRKSM 540	
DB	534 m m k y t l v t g t g n a n l a f l m p q a g t k g t s n y n d e l e n i h t s q f p a p b e p f a g y r r k s m 593	
OY	541 AWMTCYSNRLPFLVONGTLTVAAKYVRSMWTYLSEGSNPEDMMIPGLVRNCFEYRNAR 600	
DB	594 a w m t c y s n r l p f l v o n g t l v a a k y v r s m w t y l s e g s n p e d m m i p g l v r n c f e y r n a r 653	
OY	601 STWNSPAQOPPESTESSSSSDSSTSGSSTTPSTNNSTTNPNANNTOQSNTPPDOQN 660	

Db 654 swssppqppstessssdstsgsstpstnstntnnpnntqgsntlrpdpqgn 713
QY 661 POPAQP 666
| | | | |
Db 714 pqpapq 719

RESULT 4
AAV56106
ID AAV56106 standard; Protein; 420 AA.
XX
AC AAV56106;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1a TER isolate a) protein sequence.
XX
KM Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1a;
KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis.
XX
OS Streptococcus pneumoniae.
XX
PN 2A9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UYWI-) UNIT WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
DR WPI; 1999-601770/51.
DR N-PSDB; AA235939.
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
PS Claim 11; Fig 4; 63pp; English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC S. pneumoniae can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate protein
CC sequence from the present invention.
XX
SQ Sequence 420 AA;

Query Match 63.4%; Score 2208; DB 20; Length 420;
Best Local Similarity 99.0%; Pred. No. 6, 6e-146;
Matches 416; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 192 EOXEKAVNPTITGLOSLKASNSYPAYMDNYLKEVINQVEEETGYNLTITGMQVYTVWDQ 251
| | | | |
Db 1 egqykavntpdtqglqslksasnypaymdnylkevinqveeetgynllitgmqvytlrvdq 60

QY 252 EAQKHLMDIYNDEYVAYPDELOAVSTIVDVSNKGVIQAOLGARHOSNSVFGINCAVET 311
| | | | |
Db 61 eaqkhlwdiynntdeyaypdpdelqvastlvdvsngkvlaqlgarhgsnvsfginqavet 120
QY 312 NRDMGSTMKPTTDYAPALAEYGVSDSTATIVHDEPYNPGTNTPYVNDRCYFPGTITQYA 371
| | | | |
Db 121 nrwsgtmkpttdyapalegyvestatlvhdepyngtntpyndrcyfgnltlqya 180
QY 372 LOOSRNPVAVETLNKVLNRAKTFNLGLIDYPSIHSNAISSNTBESDKRYGASSEKMA 431
| | | | |
Db 181 lqsgsnvpavetlnkvlnraktfnlglidypsihsnaisntlesdkkygassekma 240
QY 432 AAYAAFPANGSTYKPKYIHKVVSDDGSEKESVNGTRAMKETTYAMNTDMAKTYLYITGTG 491
| | | | |
Db 241 aayaaafangstlykpmiyihkvvfdsgekefsnvgtramketltaymntdmkltvisygtg 300
QY 492 RNMYLAMLPGAGKTGTSNYNDEFEENIKTSOPAPDELFAGYTRKYSMAVWGYSRLT 551
| | | | |
Db 301 rmaylawlpagktgtsnyndeeleniktsqvapdeliaqylrkysmavwvqysnrlt 360
QY 552 PLVGNGLTVAAKYRSMWMTYLSGSDNPDNPIEGLYRNGCEFEVFKNGARSTWSPAPQP 611
| | | | |
Db 361 plvngnltvaaakyrsnmtylsgsdnpedwnipeglyrngefivfkngarstwsppqp 420

RESULT 5
AAU35135
ID AAU35135 standard; Protein; 778 AA.
XX
AC AAU35135;
XX
DT 13-FEB-2002 (first entry)
XX
DE Enterococcus faecalis cellular proliferation protein #422.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Enterococcus faecalis.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS52994.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 10728; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhii, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The


```

Db      245  tgmdivlnvdeaqkhlwdiynldeyaypddelqvastivdsngkviaqigarthgsn 304
QY      301  VSEGINQAVETNRDMG 316
Db      305  vsfgingaveucnrwq 320

RESULT 7
AAU03646
ID      AAU03646 standard; Protein: 462 AA.
XX
AC      AAU03646;
XX
DT      12-SEP-2001 (first entry)
XX
DE      Group B Streptococcus antigenic protein, ID-122.
XX
KW      Group B Streptococcus; encapsulated bacterium; therapeutic; sepsis;
KW      meningitis; neonate; antigenic; vaccine; infection; genital tract;
KW      capsid polysaccharide vaccination.
XX
OS      Streptococcus agalactiae.
XX
PN      WO200132882-A2.
XX
PD      10-MAY-2001.
XX
PF      07-SEP-2000; 2000MO-GB03437.
XX
PR      07-SEP-1999; 99GB-0021125.
XX
PA      (MICR-) MICROBIAL TECHNIQS LTD.
XX
PI      Le Page RWF, Wells JM, Hanniffy SB;
DR      WPI; 2001-316444/33.
DR      N-PSDB; AAS07063.
XX
PT      New polypeptides derived from Streptococcus agalactiae are useful to
PT      provide detection of, and vaccination against, Group B Streptococcus
PT      infections, particularly to prevent infection in neonates -
XX
XX
XX      Claim 1: Fig 1; 178pp; English.
XX
AAU03601-AAU03722 represent Group B Streptococcus (Streptococcus
CC      agalactiae) amino acid sequences of the invention. S. agalactiae is an
CC      encapsulated bacterium which is a major pathogen of humans causing sepsis
CC      and meningitis in neonates as well as adults. The S. agalactiae antigenic
CC      polypeptides are used to vaccinate against Group B Streptococcus
CC      infections, particularly to prevent infection in new born children
CC      arising from the maternal genital tract. An immunogenic composition is
CC      useful in the preparation of a medicament for the treatment of
CC      propyriaxis of Group B Streptococcus infection. The invention does not
CC      have the disadvantages of varied response rate associated with prior art
CC      capsid polysaccharide vaccination against Group B Streptococcus.
XX
XX
SQ      Sequence 462 AA:

Query Match      39.18; Score 1361; DB 22; Length 462;
Best Local Similarity 62.0%; Pred. No. 1e-86;
Matches 251; Conservative 73; Mismatches 81; Indels 0; Gaps 0;

QY      2  IYNNKKNLIDGSESRVNNQANDIPTDYKATYSIEDHREFDHRGIDTIRILGAFILRLN 61
Db      58  vvgngnslidlgsekresvadsdipnlvnaitsadkxfikrnydyirllgaahnl 117
QY      62  QSNSIAGSTLTQOLKILYFSTSTSDQTSRKAQAMLAIOLEPOKAKOEILTYINKV 121
Db      118  vsnntggstldqglkllyfstnksqgltkrksqewlalgmerkytkceillfylnkv 177
QY      122  YMSGNGYGMQTAQNYVYGGKDLNLISLPOLALLACGMPQAPNOYDFYSHPEAAQDRRLVLS 181

```

```

Db      178  ymgngnygmrttakysfygkdlkelstlaqallaglpqptqdydknpesqtrrntvlq 237
QY      182  EMKNQGYISAEQYKAVNPTITDGLQSLKSASNPAWMNDNLKEVINQVEETGYNLTT 241
Db      238  qmygdkniskkeydqavaltgdglkelqkkslypkymdnylkqvisvkkqtkgdlfta 297
QY      242  GMDVYTNVNDQEAQKHLMDIYNTDEVVAYPPDELOVASTIVVNSGKYIAQLGARHSSNV 301
Db      298  glkvytnlndaqqlqdydinsdylaypnmelqastlmdatnqkilaqldghnqeni 357
QY      302  SFGINQAVETNRDMGSTMKPTITDYAPALEYGVYDSTATIYHDEPYNPGTNPYVNMNRG 361
Db      358  sfngsgvltddqwtgsktmkplisayapaldsgvynstgslndsvyypgtatqldwdrg 417
QY      362  YFGNITLOVALQOSRNPVAVETLNKVGINRAKTFELNGCIDPSI 406
Db      418  ymgwmnqtaiqgsrnpvavraleaagldcaakfleklylgyppem 462

RESULT 8
AAU34286
ID      AAU34286 standard; Protein: 727 AA.
XX
AC      AAU34286;
XX
DT      14-FEB-2002 (first entry)
XX
DE      Staphylococcus aureus cellular proliferation protein #562.
XX
KW      Antisense; prokaryotic cellular proliferation protein;
KW      antibiotic; antibacterial; drug design.
XX
OS      Staphylococcus aureus.
XX
PN      WO200170955-A2.
XX
PD      27-SEP-2001.
XX
PF      21-MAR-2001; 2001MO-US09180.
XX
PR      21-MAR-2000; 2000US-191078P.
XX
PR      23-MAY-2000; 2000US-206848P.
XX
PR      26-MAY-2000; 2000US-207727P.
XX
PR      23-OCT-2000; 2000US-242578P.
XX
PR      27-NOV-2000; 2000US-253625P.
XX
PR      22-DEC-2000; 2000US-257931P.
XX
PR      16-FEB-2001; 2001US-269308P.
XX
PA      (ELIT-) ELITRA PHARM INC.
XX
PI      Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI      Yamamoto RT, Xu HH;
XX
DR      WPI; 2001-611495/70.
DR      N-PSDB; AAS52145.
XX
PT      New polynucleotides for the identification and development of
PT      antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
XX      Example 3; Seq ID No 5782; 511pp; English.
XX
PS      The invention relates to antisense inhibitors of genes essential to
PS      prokaryotic cellular proliferation, their use in identifying the
CC      genes, their use in the discovery of novel antibiotics, the essential
CC      genes themselves and the encoded proteins. The prokaryotes used are
CC      Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC      pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC      invention is also useful for the identification of potential new targets
CC      for antibiotic development. The antisense nucleic acids can also be used
CC      to identify proteins used in proliferation, to express these proteins,
CC      and to obtain antibodies capable of binding to the expressed proteins.
CC      The proteins can be used to screen compounds in rational drug discovery

```


THIS PAGE BLANK (USPTO)


```

Db 529 PIALROPNGVWSPQNDRRYSSEGRVMLVDALTRSNVPTNLGMLGLPAVETWIKL 588
QY 400 GIDYPIHISNAT---SSNTTESDKKYGASSEKMAAAYAFANGCTYYKMYIHKVYFSD 456
Db 589 GVPKDLHPVPAALGALNTPIE-----VAAQFOTIAGGNAPLSALRSVIAED 639
QY 457 GS--EREFSNVGIRAKETAYAMTDMKVLTYGTGR--NAVIALPQAGKTGTSNYTD 512
Db 640 GKLYGSFPO-AERAVPAQAAYILTWMTMOGVGRTGRCRGAKAYPNLHLAKGTGTIN--- 695
QY 513 EELENNIKTSQYAPDELPAGTTRKYSMAVWIGYSNRLLPVLGNGLTVAARYRSMYTL 572
Db 696 ---NNV-----DTWFAIGDSTVTTITWGRDNN-OPTKLYGASGMSIYQ---RYL 739
QY 573 SEGSDNEWNI--PE-----GLYRNGEPVFKNGAR--STWNSPAPQOPSTESSSSSDS 623
Db 740 A-NQTTPLVLPEDPIADMGVDYDGNFVCGSGMRLLPVWTS-----DPOSJCQOS 789
QY 624 STSQSSSTPTSTNNSTTNPNNNTTQOQNTTPOQONPOPAQ 665
Db 790 EMQO-----QPSGNPFOSSQPOQPOQ-QPAQ 816

RESULT 12
US-08-245-511-24
; Sequence 24, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6

```

```

; IMMEDIATE SOURCE:
; CLONE: SP042
; US-08-245-511-24

Query Match 11.7%; Score 406; DB 2; Length 77;
Best local Similarity 100.0%; Pred. No. 6.2e-24;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 TTGMGYTNTVNDQAKHMDIYITDEVAVPDELOVASTIVVSGKYAQLGARHQS 299
Db 1 TTGMGYTNTVNDQAKHMDIYITDEVAVPDELOVASTIVVSGKYAQLGARHQS 60
QY 300 NVSEGINQAVETNRDNG 316
Db 61 NVSEGINQAVETNRDNG 77

RESULT 13
US-08-600-993A-24
; Sequence 24, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:

```

```

: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
: IMMEDIATE SOURCE:
: CLONE: PARC0468 PBP 1B OQLL
US-08-481-435-8

```

```

Query Match      14.7%; Score 512; DB 3; Length 844;
Best Local Similarity 28.6%; Pred. No. 2.5e-30;
Matches 201; Conservative 99; Mismatches 275; Indels 128; Gaps 30;

```

```

QY 9 LIADGSRRRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRNLSQ 67
DB 204 MISSPNGEQRLEVPKSGPDLIVDTLLATEDRHFEHDSISLISIGRAVLNLTAGRTVQ 263
QY 68 GGSFTLQOLIKLTYFSTSDQTSRKAEAMLAIOLEKATKOEILTYINKVYM---- 123
DB 264 GASTLTLLVNLFLS---SERSYWRKANEAVALIMDARSKDILLELYNMEVYLGGSG 320
QY 124 SNGYGMOTAAQNYGKDLNLSLPQLALLAGMPQAPNOQYPYSHPEAAQDRNLVISEM 183
DB 321 DNEIRGFPLASLYYFGRFVEEISLDQALLVGMVKGASTYNDWPNPKLALERRNLVLRLL 380
QY 184 KNOGYSAEQYKAVNPTITDGLSLKSA-SNYPAYMDNYLKEVI---NOVEETGYNL 238
DB 381 QOOQIITDELVDMLSARPL--GVOPRGVISPQAPFMQLVROELQAKLDKKKLSGVKI 438
QY 239 LTTGMDVYTNVDOEA-----OKHMDIYNTDEVYVAPDDELQVASTIVDS 284
DB 439 FTT----FDSVAQDAEKAKEAGIGIPALKKQRLSD-----LETAIVVDPRF 480
QY 285 NGKYIAQLGARHQSNSVF-GINQAVETNRDMGSTMKPTIDYAPALEXYGVYDSTATIVHD 343
DB 481 SGEVYAMVG---GSEPOFAGYNRAMQARRSIGSLAKPAT-YLTALSQPKIYRLNTWIAD 535
QY 344 ERYNPGTNTPYV---NMDRGY--FGNTILOVALOOSRNVPAVELLNKVGAKRTFPLNG 398
DB 536 APIALROPNGOWSPONDNRYSSEGRVALVDALFRSMNVLPVNLGMLGLPAVETWIK 595
QY 399 LGIDYPSIHYENAI---SSNTTESDKKYGASEKMAAAYAAAFANGCTYKKPYIHKVYFS 455
DB 596 LGVPRDQDHPYRAMLGLNLPLIE-----VAQAFTIASSGNAAPLSALRSVIAE 646
QY 456 DGS--EKEFSNVGTRAMKETTAYMMTDMKTYLYLYGTGR--NAYLAWLPQAKTGTSNYT 511
DB 647 DGKVLVQSFPQ--AERAVYPAQAAAYLTLMTMOOVYVGRGTROLGAKYRPNLHLAGKTGTN-- 703
QY 512 DEETIENKIKTSOFVAPDLELFGYTRKYSMAVWTGYSNRLTPLYGSLVAAKVVYSMTY 571
DB 704 -----NNV-----DTWPAIDGISTVTYITWVGSDNN-QPFTKLYGASGAMSTYO--RY 746
QY 572 LSEGSNPEDWNI--PE-----GLYNGEVEFKNGAR--STWNSPAPQOPPTRESSSSSD 622
DB 747 LA-NQTPPLMLVPRPEDLADMGVDYDGNFVCSGGRILPLVMTS-----DPSLCOQ 796
QY 623 SSTSOSSSTTPSTNNSTTNPNNNTQOSNTTTPDQONONPQPAQ 665
DB 797 SEMOQ-----QPSGNPFQSSSQPOQOPQO-QPQAQ 824

```

```

RESULT 11
US-08-481-435-9
: Sequence 9, Application US/08481435
: Patent No. 6027906
: GENERAL INFORMATION:
: APPLICANT: Balganes, Tanjore S
: APPLICANT: Town, Christine
: TITLE OF INVENTION: No. 6027906el Polypeptides
: NUMBER OF SEQUENCES: 42

```

```

: CORRESPONDENCE ADDRESS:
: ADDRESSEE: White & Case
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2787
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/481,435
: FILING DATE: 10-JUL-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: IN 580/MAS/94
: FILING DATE: 01-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: SE 9404072-2
: FILING DATE: 24-NOV-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Sterner, Richard J.
: REGISTRATION NUMBER: 35,372
: REFERENCE/DOCKET NUMBER: 1103326-151
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 819-8783
: TELEFAX: (212) 354-8113
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 836 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
: IMMEDIATE SOURCE:
: CLONE: PARC0469 PBP 1B del 8
US-08-481-435-9

```

```

Query Match      14.0%; Score 489; DB 3; Length 836;
Best Local Similarity 28.1%; Pred. No. 1.3e-28;
Matches 197; Conservative 97; Mismatches 274; Indels 134; Gaps 30;

```

```

QY 9 LIADGSRRRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRNLSQ 68
DB 204 MISSPNGEQRLEVPKSGPDLIVDTLLATEDRHFEHDSISLISIGRAVLNLTAGR--- 260
QY 69 GSTLQOLIKLTYFSTSDQTSRKAEAMLAIOLEKATKOEILTYINKVYM----S 124
DB 261 ----TVQVLKLNFLS---SERSYWRKANEAVALIMDARYSKDRILLELYNMEVYLGGSD 313
QY 125 NGNYGMOTAAQNYGKDLNLSLPQLALLAGMPQAPNOQYPYSHPEAAQDRNLVISEM 184
DB 314 NEIRGFPLASLYYFGRFVEEISLDQALLVGMVKGASTYNDWPNPKLALERRNLVLRLLQ 373
QY 185 NQGYISAEQYKAVNPTITDGLSLKSA-SNYPAYMDNYLKEVI---NOVEETGYNL 239
DB 374 QOOQIITDELVDMLSARPL--GVOPRGVISPQAPFMQLVROELQAKLDKVDLSGVKIF 431
QY 240 TTGMDVYTNVDOEA-----OKHMDIYNTDEVYVAPDDELQVASTIVDSN 285
DB 432 TT----FDSVADDAEKAKEAGIGIPALKKQRLSD-----LETAIVVDPRF 473
QY 286 GKVIQAOLGARHQSNSVF-GINQAVETNRDMGSTMKPTIDYAPALEXYGVYDSTATIVHD 344
DB 474 GEVRLAMVG---GSEPOFAGYNRAMQARRSIGSLAKPAT-YLTALSQPKIYRLNTWIAD 528
QY 345 PYNPGTNTPYV---NMDRGY--FGNTILOVALOOSRNVPAVELLNKVGAKRTFPLNG 399

```

```

RESULT 9
US-08-481-435-7
; Sequence 7, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Steiner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 844 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; IMMEDIATE SOURCE:
; CLONE: PARC0438 PBP 1B QOAA
US-08-481-435-7

Query Match 14.8%; Score 514; DB 3; Length 844;
Best Local Similarity 28.6%; Pred. No. 1.8e-30;
Matches 201; Conservative 99; Mismatches 275; Indels 128; Gaps 30;

QY 9 LTAADGSEBRYVNAQANDPTDYLKAVISTEDHFFPDHGRGIDTIRIIGAFILRNQS-NSIQ 67
Db 204 MISSPGEORLVEPRSGFDLVDTLATEDRHFYEHGDISLXISGRAVLANTTAAGRTQ 263
QY 68 GSGTLTQOLIKLTYFSTSDOTISRKAOEAWLAIOLEOKATKOELLTYINKVY---- 123
Db 264 GASTLTAALVKNFLS---SERSYWRKANAYVALLIMDARYSKDRILLELMNVIYIGSG 320
QY 124 SNGNGMGTAAONYGKDLNNISLPOLALLAGMPQAPNOYDPSHPEAAODRNLVLSM 183
Db 321 DNEIRGFPLASYFEGRPEVELSLDQALLVGWVGASITYNPRNPKLALERRNLVRL 380
QY 184 KNGYISAQYEAQVPTPTDGLQSLKSA-SNPRATMDNYLKEV-----NOVEEELGYML 238
Db 381 OOOOIIDOLYDMLSLARPL--GVQPRGVVISPPQAPFMOLVRELQAKLDKDYKSLGVKI 438
QY 239 LTTGMDEVYTNVQEA-----QKHLMDIYNDEYVAYVDDDELQVASTIVDS 284

```

```

Db 439 FTT-----FDSVADAAEKAKEGIPALKKOKRLSD-----LETAIVVDRF 480
QY 285 NGKVIQOLGARHOSNVSF-GIQNAVEETNMDKSTMKPTTDYAPALEYGYDSTATIVHD 343
Db 481 SGYERAMVG---GSEPOFAGYNRAQARRSIGSLAKPAT-YLTALSQPKYIRLNTWIAD 535
QY 344 EPNYRPGTNTPPVY---NWDGKY--FGNITLYVLOOSRNPVPAVETLNKVLNRAKTFLENG 398
Db 536 APLALQPNQVWSPONDDRRYSESGRYMLVDLITSMNVPYTNLGMALCLPAVTEWIK 595
QY 399 LGIDYPSIHYSNAI---SSNTTESDKKYGASSERMAAAYAFANGGTYYKPMYIHKVFS 455
Db 596 LGVPKQOLHPVPMALGALNLFTPE-----VAQAFQTIASGGRAPLALSRIYAE 646
QY 456 DGS--EKESNVGTRAMKETTAATMDAKTUTLYTGCR--NAYLMLPQAGCTGTSNNT 511
Db 647 DGKVLQSFPO-AERAVPAQAAVLTMTMOQVVOGRGROLGAKYPNLHLAGRTGTTN-- 703
QY 512 DELEENHKTSQFVAPDELPAQYTRKYSMAVMTGYSNRLPLVGNGLTVAKAYRSMTY 571
Db 704 -----NNV-----DTMFAGIDGSTVITTWVGRDNN-OPTKLYGASGAMSTIYQ---RY 746
QY 572 LSEGSNPEDMNT--PE-----GLYRGEFVFKNGAR--STWNSPAPQQPSTESSSSSD 622
Db 747 LA-NQTPPLNLVPEPDIDMGVDYDGNFVCSGMRILPWTS-----DPQSLCQ 796
QY 623 SSTSSQSSSTPTSTNNSTTRPNNNNTQOQSNTPPQOQNOQPAQ 665
Db 797 SEMQO-----QPSGNPQDSSQPOQOQPOQ-QPAQ 824

RESULT 10
US-08-481-435-8
; Sequence 8, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Steiner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 844 amino acids

```


TELEFAX: (212) 354-8113
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 828 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-481-435-2

Query Match 16.3%; Score 567; DB 3; Length 828;
 Best Local Similarity 23.4%; Pred. No. 1.7e-34;
 Matches 202; Conservative 129; Mismatches 244; Indels 290; Gaps 30;

```

QY 1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRIL-GAFLR 59
DB 27 QIYSADGELIAOYGERKRRIPVTLDPPEMVKAFIATEDSREYEHNGVDPVGIFFRAASYA 86
QY 60 NIQNSLGGSTLTQOLIKLTFYSTSTSDQTSRKAQEMALIQLEOKATKQELITYYIN 119
DB 87 LFSGASHQCASTITQOLANFPLS---PERTLMRKIKEVFLAIRIQLLTKDEILELYLN 143
QY 120 KYVMSNGNYGMOTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAAQDRRLV 179
DB 144 KIYLGIRAYGVGAAQVYFGKTVDOQLTNEMAVIAGLPKAPSTFNFLYSMDRAVARRNV 203
QY 180 LSEMRKOGYISAEOYEKAVNTPITDGLQSLKSASNPAYMDNLYKEVINQVEE---ETGY 236
DB 204 LSRMDEGIYTOOQFDQTEAIANNYHAPETAFSAFYLSENVROMYNRGESAVEDGY 263
QY 237 NLTGMDVYTVNDQEAQ-----KH-----LMDIYN-----TDEEYA 268
DB 264 RIYTT---ITRRVQQAQAQAVRNINVDYDMRHCRGPANVLMKVGSAANDNKITPTLKA 320
QY 269 YP-----DDEL 274
DB 321 LPTYPGLPAAVTSANPQATAMLADGSTVALSMEGVRWARPYSRDTQOGPTPRKVTYDVL 380
QY 275 Q-----VASIIVDVS--NGKVINQLGA--RHQSSNSFGLNOAV 309
DB 381 QNGQOIWVROVDAMWMLAOVPENSALVISINPQNGAVMLVGGFDENQSK---FNRAIT 435
QY 310 ETNRDMGSTMKPIITDYPALAEYGVYDSTATIVHDEPYNPGTVPYMNDRG----- 361
DB 436 QALROVGSIKRPL--YTALMDKGL--TLASMLND-----VPSLRDASAGSMWRK 483
QY 362 ----YFGNITLQVALQGSRNVPAYETLTKVGLNRAKTFELNGIDIDYPSIHVSNAISNT 416
DB 484 NSPPQYAGPIRLRGLGQSKNVMYVAMRAMGVDAEYLQRFGRFAONIVHTESLA--- 540
QY 417 TESDKKYGASS---EKMAAAYAFAFGNGTYYKPMYIHKKVFESDG----- 457
DB 541 ----LGSASFTPMQOVANGVAVMANGFLVDPMWFLISKIENDGGVIFPAKPRVACPECD 594
QY 458 ----SEKEFSNV-----GTRAMKETTAYM--MTDMKMTVLVLYGRRNAY----- 495
DB 595 IYVYIGDQTKSVNLNENVDYEAIVSREQQNSVPMFQLOANQALAKGAQDEYADPHVIN 654
QY 496 --LAWLPQA-----GKTGSTNNTDEIEIENHIKTSOF 524
DB 655 TPLAFLIKALMTNIFGEFGMGQGTWRAGRDQLQRDIGKKTGTN----- 699
QY 525 VAPDELFCAGTYTKYSMAVVTGYSNRLTPLYVNGCLYVAAVYSSMMTYLSEGSNPEDWNIP 584
DB 700 SKSDAMFSGYGVGVYTSVMIGFDDHRRNL--GHTTASGAIDQISGYEGCA----- 748
QY 585 ECLYRNGEYVFVNGARSTNS-----PAPOP--PTSESSSSSDSTSSSSTPTST 635
DB 749 -----KSAQAPAMDAYMKAVLEGVPRQPLTPPGIYVINDNSTGQ----- 788
QY 636 NNSITTNPNNTQQ---SWTTPDQ 657
DB 789 ----LANGSNSREYFIETGPTQQ 809

```

RESULT 7

US-08-481-435-10

Sequence 10, Application US/08481435

Patent No. 6027906

GENERAL INFORMATION:

APPLICANT: Balganes, Tanjore S

APPLICANT: Town, Christine

TITLE OF INVENTION: No. 6027906el polypeptides

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESS: White & Case

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,435

FILING DATE: 10-JUL-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IN 580/MAS/94

FILING DATE: 01-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: SE 9404072-2

FILING DATE: 24-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Steiner, Richard J.

REGISTRATION NUMBER: 35,372

REFERENCE/DOCKET NUMBER: 1103326-151

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 819-8783

TELEFAX: (212) 354-8113

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 850 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Escherichia coli

IMMEDIATE SOURCE:

CLONE: PARC0571 PBP 1A Q0AA

US-08-481-435-10

Query Match 15.9%; Score 555; DB 3; Length 850;
 Best Local Similarity 23.1%; Pred. No. 1.4e-33;
 Matches 200; Conservative 129; Mismatches 246; Indels 290; Gaps 30;

```

QY 1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRIL-GAFLR 59
DB 49 QIYSADGELIAOYGERKRRIPVTLDPPEMVKAFIATEDSREYEHNGVDPVGIFFRAASYA 108
QY 60 NIQNSLGGSTLTQOLIKLTFYSTSTSDQTSRKAQEMALIQLEOKATKQELITYYIN 119
DB 109 LFSGASHQCASTITQOLANFPLS---PERTLMRKIKEVFLAIRIQLLTKDEILELYLN 165
QY 120 KYVMSNGNYGMOTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAAQDRRLV 179
DB 166 KIYLGIRAYGVGAAQVYFGKTVDOQLTNEMAVIAGLPKAPSTFNFLYSMDRAVARRNV 225
QY 180 LSEMRKOGYISAEOYEKAVNTPITDGLQSLKSASNPAYMDNLYKEVINQVEE---ETGY 236
DB 226 LSRMDEGIYTOOQFDQTEAIANNYHAPETAFSAFYLSENVROMYNRGESAVEDGY 285

```


Db 185 SEMKNOGYSABOYERKAVNPTDGLQSLKSASVPAVMNYKEVINQVEETGYMLLT 244
QY 241 TGMVYTNVDOEAKHLMIDYNTDEYVAPDPDELQVASTIVDSNGKYIAQLGARHOSSN 300
Db 245 TGMVYTNVDOEAKHLMIDYNTDEYVAPDPDELQVASTIVDSNGKYIAQLGARHOSSN 304
QY 301 VSEGINQAVETNRDMG 316
Db 305 VSEGINQAVETNRDMG 320

RESULT 5
US-08-731-716-2
; Sequence 2, Application US/08731716
; Patent No. 5789202
; GENERAL INFORMATION:
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Rokey, Pamela K.
; APPLICANT: Zhao, Genshi
; APPLICANT: Rostock, Paul R. Jr.
; APPLICANT: No. 5789202ris, Franklin H.
; TITLE OF INVENTION: Penicillin Binding Protein From
; TITLE OF INVENTION: Streptococcus Pneumoniae
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,716
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-10,867
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-731-716-2

Query Match 18.3%; Score 636; DB 1; Length 731;
Best Local Similarity 27.7%; Pred. No. 8.2e-40;
Matches 19; Conservative 108; Mismatches 252; Indels 142; Gaps 19;

QY 2 IYDNKNLIADLGSERRVNAQANDIPTDLVKAIVSIDHREFDRGIDTIRILGAPLRNL 61
Db 95 IYDREKEAGALSGOKTYVELTDISKLNQNAVATIEDRSFYKNDGINYGRF--FLAIY 151
QY 62 QSNLSGGSTITTOOLANNAVLS--QDQYERKAKFEFLALELSKYSKEQILITMYLNA 208
Db 152 TAGRSGGSGTITTOOLANNAVLS--QDQYERKAKFEFLALELSKYSKEQILITMYLNA 208
QY 122 YMSNGNVMOTAAQNYGKXUNLSTPOLALLAGMPQAQYQDYSHPEAAODRRNLVLS 181
Db 209 YFNGNVMGVEDAKKRYGVASAEVSLQAAATLACMLKGPFLYNPLANSVEDSTNRDYLQ 268
QY 182 EMKNOGYSABOYERKAVNPTDGLQSLKSASVPAVMNYKEVINQVEET 234

Db 269 NMVAGYIDKNOETEEAEVDMTSQLHDKYECKISDYRPSYFADAVNVAASKYMLTEEE- 327
QY 235 GYNLTITGMVYTNVDOEAKHLMIDY-NTDEYVAPDDEL-QVASTIVDSNGKYIAQL 292
Db 328 --IVNNGYRIYTELDQNYQANMOIYENFSLFRAEDGTFFAQSGLALEKTCGVNGV 384
QY 293 GAROSSNVSP-GINQAVETNRDMGSMKPTTPDAPALEXGVYDSTATIYVDEPY----- 346
Db 385 QGVADNDKCTGFNRNFNYATQSKRSPGSTITPLVYTPPAVEGMALNKQLDNHTKQYDSTYK 444
QY 347 -NYPGINT---PYVNDRGYFGNITLQYALQOSRNVPAVETLNKVGILNRAKTFPLNGLGI 401
Db 445 DNYAGIKTSREVPMYQ-----SLAESLNPAAVATVDLGVDKA----- 482
QY 402 DYPSIHNSMAISSNTTESDKKYGASSEK-----MAAAYAFANGGT 442
Db 483 -----FAGEKEFGILMERKVDRLGVALGSGVETNPLOMAQAYAFANEGL 527
QY 443 YKPMYTHKVFSDGSE-KEFSNVTGRANKETTYMMTDMKTYLTGTGRNAVYLAFLPQ 501
Db 528 MPEHFTSRLENASGOYIASHKNSQKRVIDKSVADKMTSMLGTFGTGTGSSPADYVM 587
QY 502 AKGTGTSNYTDEIEIENIKTSQYVAPDELPAGYTRKYSMAVMTGY-----SNRLTPLVN 556
Db 588 AKGTGTT---EAVFNEBYTS---DQWVIGYTPDVVISHWLGFPTTDEHNHYLAGSTSN 638
QY 557 CLTYAAKVYSMMTYLSEGSNPEDMWNPBGLIRNGEEVFNKGARSTVNSPAPQPPETES 616
Db 639 G--AAHFNRNIANTIL-----PYTPG 657
QY 617 SSSSDSSTSQSSSTPSTNNSTTNPNNNTQOS 650
Db 658 STTVENAYKONGIAIPANTKRQVOTNDNSQTDN 691

RESULT 6
US-08-481-435-2
; Sequence 2, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Steiner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783

```
RESULT 3
US-08-245-511-4
; Sequence 4, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 343-1684
; TELEFAX: 201 487-5800
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-245-511-4

Query Match          46.5%; Score 1621; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 6.7e-115;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 305 VSFGINQAVETNRDNG 320
|||||
RESULT 4
US-08-600-993A-4
; Sequence 4, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 343-1684
; TELEFAX: 201 487-5800
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-600-993A-4

Query Match          46.5%; Score 1621; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 6.7e-115;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Db 1 KIYDNKNQLIADLGSERRVANAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
QY LOSNSLOGGSTLTQOLIKLIFYESTSTSDQTSRKAQEAAMLAIQLEKATKOEILTYINK 120
Db 61 LOSNSLOGGSTLTQOLIKLIFYESTSTSDQTSRKAQEAAMLAIQLEKATKOEILTYINK 120
QY 121 VYNSNGNYGQTAQNYGYGDLNNLSLPOLALLAGMPQAPNOQDPYSHPEAODRRLVL 180
Db 121 VYNSNGNYGQTAQNYGYGDLNNLSLPOLALLAGMPQAPNOQDPYSHPEAODRRLVL 180
QY 181 SEMKNQGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYMLT 240
Db 181 SEMKNQGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYMLT 240
QY 241 TGMQVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDVSNQKVIQOLGARHOSSN 300
Db 241 TGMQVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDVSNQKVIQOLGARHOSSN 300
QY 301 VSEGINQAVETNRDMSGTMRPTDYAPALEGYDYSTATIVHDEPNYPGTNTPVYMWDR 360
Db 301 VSEGINQAVETNRDMSGTMRPTDYAPALEGYDYSTATIVHDEPNYPGTNTPVYMWDR 360
QY 361 GYFGNITLQYALQOOSRNVPAVETLNKGLNRAKTFNLGLGIDPYSIHYSNAISSNTTESD 420
Db 361 GYFGNITLQYALQOOSRNVPAVETLNKGLNRAKTFNLGLGIDPYSIHYSNAISSNTTESD 420
QY 421 KKYGASSEKMAAAYAAFAANGTYKKPMYIHKVYFSDSEKEFSNVGTRAKETAYAMTD 480
Db 421 KKYGASSEKMAAAYAAFAANGTYKKPMYIHKVYFSDSEKEFSNVGTRAKETAYAMTD 480
QY 481 MKKTVLTYGGRNAYILAMLPQAGKTGTSNTYDEIEHNIKTQFVADDELFACTTRKYSM 540
Db 481 MKKTVLTYGGRNAYILAMLPQAGKTGTSNTYDEIEHNIKTQFVADDELFACTTRKYSM 540
QY 541 AVMTGYSNRLLTPLVGNGLTYAAKYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
Db 541 AVMTGYSNRLLTPLVGNGLTYAAKYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
QY 601 STWNSPAPQOPPTSTESSSSSDSTSSQSSSTPTNNSTTNPNNNTQOQNTTTPDOONON 660
Db 601 STWNSPAPQOPPTSTESSSSSDSTSSQSSSTPTNNSTTNPNNNTQOQNTTTPDOONON 660
QY 661 POPAOP 666
Db 661 POPAOP 666

RESULT 2
US-08-481-435-6
Sequence 6, Application US/08481435
Patent No. 6027906

GENERAL INFORMATION:
APPLICANT: Balganes, Tanjore S
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 99.5%; Score 3466; DB 3; Length 682;
Best Local Similarity 99.4%; Pred. No. 5,4e-254;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNQLIADLGSERRVANAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
Db 17 KIYDNKNQLIADLGSERRVANAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 76
QY 61 LOSNSLOGGSTLTQOLIKLIFYESTSTSDQTSRKAQEAAMLAIQLEKATKOEILTYINK 120
Db 77 LOSNSLOGGSTLTQOLIKLIFYESTSTSDQTSRKAQEAAMLAIQLEKATKOEILTYINK 136
QY 121 VYNSNGNYGQTAQNYGYGDLNNLSLPOLALLAGMPQAPNOQDPYSHPEAODRRLVL 180
Db 137 VYNSNGNYGQTAQNYGYGDLNNLSLPOLALLAGMPQAPNOQDPYSHPEAODRRLVL 196
QY 181 SEMKNQGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYMLT 240
Db 197 SEMKNQGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYMLT 256
QY 241 TGMQVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDVSNQKVIQOLGARHOSSN 300
Db 257 TGMQVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDVSNQKVIQOLGARHOSSN 316
QY 301 VSEGINQAVETNRDMSGTMRPTDYAPALEGYDYSTATIVHDEPNYPGTNTPVYMWDR 360
Db 317 VSEGINQAVETNRDMSGTMRPTDYAPALEGYDYSTATIVHDEPNYPGTNTPVYMWDR 376
QY 361 GYFGNITLQYALQOOSRNVPAVETLNKGLNRAKTFNLGLGIDPYSIHYSNAISSNTTESD 420
Db 377 GYFGNITLQYALQOOSRNVPAVETLNKGLNRAKTFNLGLGIDPYSIHYSNAISSNTTESD 436
QY 421 KKYGASSEKMAAAYAAFAANGTYKKPMYIHKVYFSDSEKEFSNVGTRAKETAYAMTD 480
Db 437 KKYGASSEKMAAAYAAFAANGTYKKPMYIHKVYFSDSEKEFSNVGTRAKETAYAMTD 496
QY 481 MKKTVLTYGGRNAYILAMLPQAGKTGTSNTYDEIEHNIKTQFVADDELFACTTRKYSM 540
Db 497 MKKTVLTYGGRNAYILAMLPQAGKTGTSNTYDEIEHNIKTQFVADDELFACTTRKYSM 556
QY 541 AVMTGYSNRLLTPLVGNGLTYAAKYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
Db 557 AVMTGYSNRLLTPLVGNGLTYAAKYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 616
QY 601 STWNSPAPQOPPTSTESSSSSDSTSSQSSSTPTNNSTTNPNNNTQOQNTTTPDOONON 660
Db 617 STWNSPAPQOPPTSTESSSSSDSTSSQSSSTPTNNSTTNPNNNTQOQNTTTPDOONON 676
QY 661 POPAOP 666
Db 677 POPAOP 682

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 03:19:04 ; Search time 32.26 Seconds

(without alignments)
504.261 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIYDNKNQIADLSESRVYN.....TOOSNTTDPQONONPQAPQ 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	666	4 US-08-961-083-2	Sequence 2, Appli
2	3466	99.5	682	3 US-08-481-435-6	Sequence 6, Appli
3	1621	46.5	320	2 US-08-245-511-4	Sequence 4, Appli
4	1621	46.5	320	2 US-08-600-993A-4	Sequence 4, Appli
5	636	18.3	731	1 US-08-731-716-2	Sequence 2, Appli
6	567	16.3	828	3 US-08-481-435-2	Sequence 2, Appli
7	555	15.9	850	3 US-08-481-435-10	Sequence 10, Appli
8	526	15.1	823	3 US-08-481-435-4	Sequence 4, Appli
9	514	14.8	844	3 US-08-481-435-7	Sequence 7, Appli
10	512	14.7	844	3 US-08-481-435-8	Sequence 8, Appli
11	489	14.0	835	3 US-08-481-435-9	Sequence 9, Appli
12	406	11.7	77	2 US-08-245-511-24	Sequence 24, Appli
13	406	11.7	77	2 US-08-600-993A-24	Sequence 24, Appli
14	370.5	10.6	532	3 US-08-481-435-12	Sequence 12, Appli
15	370.5	10.6	533	3 US-08-481-435-11	Sequence 11, Appli
16	288.5	8.3	269	2 US-08-771-716-2	Sequence 2, Appli
17	288.5	8.3	269	4 US-09-057-720A-2	Sequence 2, Appli
18	247	7.1	159	3 US-08-481-435-13	Sequence 13, Appli
19	148	4.2	1164	4 US-08-589-756-1	Sequence 1, Appli
20	148	4.2	1164	4 US-09-206-800-1	Sequence 1, Appli
21	146	4.2	2843	1 US-07-741-940-2	Sequence 2, Appli
22	146	4.2	2843	1 US-08-289-548A-2	Sequence 2, Appli
23	146	4.2	2843	1 US-08-452-654-2	Sequence 2, Appli
24	146	4.2	2843	1 US-08-370-235A-2	Sequence 2, Appli
25	141	4.0	2842	1 US-07-741-940-7	Sequence 7, Appli
26	141	4.0	2842	1 US-08-289-548A-7	Sequence 7, Appli
27	141	4.0	2842	1 US-08-289-548A-7	Sequence 7, Appli

28	141	4.0	2842	1 US-08-452-654-7	Sequence 7, Appli
29	141	4.0	2843	1 US-08-452-655B-2	Sequence 2, Appli
30	141	4.0	2843	1 US-08-452-655B-7	Sequence 7, Appli
31	141	4.0	2843	3 US-08-450-582-2	Sequence 2, Appli
32	141	4.0	2843	3 US-08-450-582-7	Sequence 7, Appli
33	141	4.0	2973	2 US-08-821-355A-7	Sequence 7, Appli
34	141	4.0	2973	2 US-09-003-687A-7	Sequence 7, Appli
35	141	4.0	2973	2 US-09-136-605-7	Sequence 7, Appli
36	139	4.0	1007	2 US-08-551-459-4	Sequence 4, Appli
37	136.5	3.9	866	2 US-08-483-101-4	Sequence 4, Appli
38	134.5	3.9	610	2 US-08-525-742-8	Sequence 8, Appli
39	133.5	3.8	478	2 US-08-456-670B-40	Sequence 40, Appli
40	133.5	3.8	933	3 US-08-293-728-2	Sequence 2, Appli
41	133.5	3.8	933	3 US-09-421-868-2	Sequence 2, Appli
42	132	3.8	1181	4 US-09-206-898-23	Sequence 23, Appli
43	130.5	3.7	1150	2 US-08-589-756-3	Sequence 3, Appli
44	130.5	3.7	1150	2 US-09-206-800-3	Sequence 3, Appli
45	130.5	3.7	1150	4 US-09-206-898-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-08-961-083-2
: Sequence 2, Application US/08961083
: Patent No. 6159469
:
: GENERAL INFORMATION:
: APPLICANT: Chol et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Stookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 666 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-961-083-2
:
: Query Match 100.0%; Score 3484; DB 4; Length 666;
: Best Local Similarity 100.0%; Pred. No. 2.3e-255;
: Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: 1 KIYDNKNQIADLSESRVNAQANDIPDLVKAISIDHRRFDRGIDTIRIICAPLRN 60
: |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

THIS PAGE BLANK (USPTO)

Mon Aug 12 08:40:00 2002

us-08-961-083-2.rag

Page 13

Db 603 ysyfmrnllaiepslkrrkfdvpglrlk--elvdkiipysspsnslp-lpkktds 654

Search completed: August 11, 2002, 04:25:30
Job time: 4356 sec

CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 727 AA:

Query Match 23.5%; Score 817.5; DB 22; Length 727;
Best Local Similarity 32.1%; Pred. No. 1,8e-48;
Matches 222; Conservative 128; Mismatches 261; Indels 81; Gaps 22;
QY 1 KIYDNKNQIADL-GSERRVNAQANDIPTDLVAIVSIEDHREFDHGIDTIRILGAFLR 59
DB 76 KIYDKNQELVKLTIDNGRHEHVNLDKVPKSMKDAVLATEDNRIFYEHGALDYKRIFFGALGK 135
QY 60 NLOSN-SLOGGSTLTQOLIKLTFSTSTSDQTSRKAQEMALAIQLEOKATKOEILYYI 118
DB 136 nlcggfsgesastltgqvvdafis---qhksglrkqaeaylsyrlqegskdrlfgyvl 192
QY 119 NKYYMNGNGYGMOTAAQNYGKDLNLSLPOLALLAGMPAQNQYDPSHPEAQRRL 178
DB 193 nkliysdgvtgikaaakyyfnkdiklnlaeeylaglpqvpnnynlydhpkaaedrknt 252
QY 179 VLSEMKNOGYISAEQYKA-----VN-TPITDGLQSLKSNYPAYMDNLTKEVIN 228
DB 253 vlylmhvkrltdkqwedakkidkanlvrtpceqrnldtngdesynvnfksslm 312
QY 229 Q--VEETGYNLTLTGADVTYNDOEAKHMDIYNTDEVYANPDELQYASTIVDSNG 286
DB 313 nkaefkdenlgnvlsgqiklylmdkdvkclqn--dvdnsgsfyknkdqvgatlldsktg 370
QY 287 KVIAGLCARQSSNVSGFIGNOAVETNRDQSTMKPTIDYAPALEGYVDSTATIVHDEPY 346
DB 371 glvalsggrfkdvvn--rrqacdpn-ptgsslkpfliaypalemmkwaohnaigdesy 427
QY 347 NPYGTNPVYMWDRGYFGNITLQYALQOASRNPVETLNK---GLNRAKTEPLNGLGID 402
DB 428 qvdgst--fnydtkshgvtvsiydalrqsfnipalkawgsvknagndapkkfaaklyln 485
QY 403 YP-SIHNSNAISSNTTSDSKKYGASSERKMAAAYAFANGSTGYKPMIHKHVNVSDESEKE 461
DB 486 yegdligpsevliggsase-----fspcqlasatafaalnggltymahsldkvvtrldgetle 539
QY 462 FSNVGRFAMKETAYWMTDMK-TVLTYGGRNAYLAWLPQAGKGTGSNTTDEIENHIK 520
DB 540 ydhtshkamadytaymlaemlkglfkpygsayghygsvgvmgaktgctygaely----- 594
QY 521 TSOFAVDEL----FAGYTRKYSMAVWTGYSNRLPPLVNGLTVAAKVRSMTWTYLSG 575
DB 595 -sgynlpdnaekdvngiftpgytmsvmgfsk-----vkgy-----g 631
QY 576 SNP---EDMNIPEGLYRNGEYFVKNGARSTWNSPAPQPPSTESSSSSDSTSGSSST 631
DB 632 ensfvshkamsdytaymlaemlkglfkpygsayghygsvgvmgaktgctygaely----- 686
QY 632 TPST---NSTTNPNNNTTQOASNTTPOQNON 660
DB 687 tnrstngsdtsanssgtaqsnmtrtsqgsrn 718

RESULT 9
AAU37184
ID AAU37184 standard; Protein; 727 AA.
XX
XX AAU37184;
XX
DT 14-FEB-2002 (first entry)
XX

DE Staphylococcus aureus cellular proliferation protein #1354.
XX
XX Antisense: prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WC200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
XX
PR 23-MAY-2000; 2000US-206848P.
XX
PR 26-MAY-2000; 2000US-207727P.
XX
PR 23-OCT-2000; 2000US-242578P.
XX
PR 27-NOV-2000; 2000US-253625P.
XX
PR 22-DEC-2000; 2000US-257931P.
XX
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX
DR N-PDB; AAS55043.
XX
PT New polynucleotides for the identification and development of
XX
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 12777; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX programmes. The antisense nucleic acid sequence is also useful to screen
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms. The present sequence represents an
XX essential prokaryotic cellular proliferation protein.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 727 AA:
Query Match 23.5%; Score 817.5; DB 22; Length 727;
Best Local Similarity 32.1%; Pred. No. 1,8e-48;
Matches 222; Conservative 128; Mismatches 261; Indels 81; Gaps 22;
QY 1 KIYDNKNQIADL-GSERRVNAQANDIPTDLVAIVSIEDHREFDHGIDTIRILGAFLR 59
DB 76 KIYDKNQELVKLTIDNGRHEHVNLDKVPKSMKDAVLATEDNRIFYEHGALDYKRIFFGALGK 135
QY 60 NLOSN-SLOGGSTLTQOLIKLTFSTSTSDQTSRKAQEMALAIQLEOKATKOEILYYI 118
DB 136 nlcggfsgesastltgqvvdafis---qhksglrkqaeaylsyrlqegskdrlfgyvl 192
QY 119 NKYYMNGNGYGMOTAAQNYGKDLNLSLPOLALLAGMPAQNQYDPSHPEAQRRL 178
DB 193 nkliysdgvtgikaaakyyfnkdiklnlaeeylaglpqvpnnynlydhpkaaedrknt 252

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 04:16:44 : Search time 214.73 Seconds
(without alignments)
1091.692 Million cell updates/sec

Title: US-08-961-083-2
Perfect score: 3484
Sequence: 1 KIYDNKNQLIADIGSERRVN.....TQOSNTTPDQONNPQAPQ 666

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 segs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCrus_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	666	19	US-09-536-784-2
2	3484	100.0	666	21	US-09-765-271-2
3	3484	100.0	666	21	US-09-765-272-2
4	3483	100.0	719	15	US-09-583-110-3973
5	3483	100.0	721	19	US-09-433-3705
6	3466	99.5	719	1	PCR-US02-03987-13423
7	3466	99.5	719	22	US-09-815-242-13423

8	3466	99.5	719	24	US-10-072-851-13423	Sequence 13423, A
9	1862.5	53.5	823	15	US-09-107-532-5667	Sequence 5667, Ap
10	1862.5	53.5	823	15	US-09-107-532A-5667	Sequence 5667, Ap
11	1751	50.3	778	1	PCR-US02-03987-10728	Sequence 10728, A
12	1751	50.3	778	22	US-09-815-242-10728	Sequence 10728, A
13	1751	50.3	778	24	US-10-072-851-10728	Sequence 10728, A
14	1751	50.3	789	15	US-09-134-000-4939	Sequence 4939, Ap
15	1621	46.5	320	5	US-08-116-541-4	Sequence 4, Appl1
16	1473.5	42.3	771	20	US-09-634-238-229	Sequence 229, App
17	1473.5	42.3	771	20	US-09-634-238-262	Sequence 262, App
18	1361	39.1	462	24	US-10-091-007-92	Sequence 92, Appl
19	817.5	23.5	727	1	PCR-US02-03987-5782	Sequence 5782, Ap
20	817.5	23.5	727	1	PCR-US02-03987-12777	Sequence 12777, A
21	817.5	23.5	727	22	US-09-815-242-5782	Sequence 5782, Ap
22	817.5	23.5	727	22	US-09-815-242-12777	Sequence 12777, A
23	817.5	23.5	727	24	US-10-072-851-5782	Sequence 5782, Ap
24	817.5	23.5	727	24	US-10-072-851-12777	Sequence 12777, A
25	815.5	23.4	730	20	US-09-611-529-4512	Sequence 4512, Ap
26	811.5	23.3	716	26	US-60-253-625-2437	Sequence 2437, Ap
27	811.5	23.3	716	26	US-60-257-931-2821	Sequence 2821, Ap
28	811.5	23.3	716	26	US-60-269-308-3840	Sequence 3840, Ap
29	809.5	23.2	746	15	US-09-134-001C-3214	Sequence 3214, Ap
30	651.5	19.0	675	20	US-09-634-238-228	Sequence 228, App
31	655	18.8	698	21	US-09-739-449-11825	Sequence 11825, A
32	655	18.8	698	22	US-09-803-110-11825	Sequence 11825, A
33	648.5	18.6	814	17	US-09-328-352-4373	Sequence 4373, Ap
34	641	18.4	834	16	US-09-252-991A-32671	Sequence 32671, A
35	636	18.3	730	15	US-09-107-433-4707	Sequence 4707, Ap
36	636	18.3	731	1	PCR-US97-19070-2	Sequence 4707, Ap
37	636	18.3	731	19	US-09-583-110-4720	Sequence 4720, Ap
38	632.5	18.2	855	17	US-09-328-352-6216	Sequence 6216, Ap
39	625	17.9	660	1	PCR-US97-05223-845	Sequence 845, App
40	625	17.9	660	1	PCR-US97-05223-876	Sequence 876, App
41	625	17.9	660	11	US-08-761-318-845	Sequence 845, App
42	625	17.9	660	11	US-08-761-318-876	Sequence 876, App
43	625	17.9	660	12	US-08-824-132-845	Sequence 845, App
44	625	17.9	660	12	US-08-824-132-876	Sequence 876, App
45	625	17.9	660	13	US-08-993-002A-5603	Sequence 5603, Ap

ALIGNMENTS

RESULT 1
US-09-536-784-2
Sequence 2, Application US/09536784
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,784
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-536-784-2

Query Match 100.0%; Score 3484; DB 19; Length 666;
Best Local Similarity 100.0%; Pred. No. 3e-289;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNLIADLGSERRVNAQANDIPDYLKAIYSIEDHREFDHRGIDTIRILGAFLRN 60
DB 1 KIYDNKNLIADLGSERRVNAQANDIPDYLKAIYSIEDHREFDHRGIDTIRILGAFLRN 60
QY 61 LOSNLSOGSSTLTQOLIKLTYFSTSDQTSRKAQEAFLAQLQKATKOEILTYYNK 120
DB 61 LOSNLSOGSSTLTQOLIKLTYFSTSDQTSRKAQEAFLAQLQKATKOEILTYYNK 120
QY 121 VYMSNGNGMGTAAQNYGKDLNNLSLPQALLAGMPQAPNOYDPSHPEAODRRNLVL 180
DB 121 VYMSNGNGMGTAAQNYGKDLNNLSLPQALLAGMPQAPNOYDPSHPEAODRRNLVL 180
QY 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNYLKEVINQVEEETGYNLT 240
DB 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNYLKEVINQVEEETGYNLT 240
QY 241 TGMDEVYTNVDEAOKHLMIDYNTDEVYVAPDELOVASTIVDSNGKVIQALGARHSSN 300
DB 241 TGMDEVYTNVDEAOKHLMIDYNTDEVYVAPDELOVASTIVDSNGKVIQALGARHSSN 300
QY 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTNPVYNNDR 360
DB 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTNPVYNNDR 360
QY 361 GFNGNITLQYALQOSRNPVAVETLNKVGILNRAKTFNLGLIDYPSIHVSNAISSNTTESD 420
DB 361 GFNGNITLQYALQOSRNPVAVETLNKVGILNRAKTFNLGLIDYPSIHVSNAISSNTTESD 420
QY 421 KRYGASSEKMAAAYAFANGGTYYPMYIHKVFSDGSEKESNVGTRAMKETTYAYMMTD 480
DB 421 KRYGASSEKMAAAYAFANGGTYYPMYIHKVFSDGSEKESNVGTRAMKETTYAYMMTD 480
QY 481 MMKTYLVGTGCHNAYLAWLPQAGKTGTSNYTDEETENHIKTSQFVAPDELFAGYRKYSM 540
DB 481 MMKTYLVGTGCHNAYLAWLPQAGKTGTSNYTDEETENHIKTSQFVAPDELFAGYRKYSM 540
QY 541 AVMTGYSNRLPLVGNGLTVAAKYRSMKTYISEGSPEDMNIPEGILRNGEFVKNGAR 600
DB 541 AVMTGYSNRLPLVGNGLTVAAKYRSMKTYISEGSPEDMNIPEGILRNGEFVKNGAR 600
QY 601 STMWSPAPQOPSTSSSSSSDSTSSSTPTNNSTTTPNNNTQOOSNTTPQOONON 660
DB 601 STMWSPAPQOPSTSSSSSSDSTSSSTPTNNSTTTPNNNTQOOSNTTPQOONON 660
QY 661 PQPAP 666
DB 661 PQPAP 666

RESULT 2
US-09-765-271-2
Sequence 2, Application US/09765271
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,271
FILING DATE: 22-Jan-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/536,784
FILING DATE: <unknown>
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-271-2

QY 1 KIYDNKNLIADLGSERRVNAQANDIPDYLKAIYSIEDHREFDHRGIDTIRILGAFLRN 60
DB 1 KIYDNKNLIADLGSERRVNAQANDIPDYLKAIYSIEDHREFDHRGIDTIRILGAFLRN 60
QY 61 LOSNLSOGSSTLTQOLIKLTYFSTSDQTSRKAQEAFLAQLQKATKOEILTYYNK 120
DB 61 LOSNLSOGSSTLTQOLIKLTYFSTSDQTSRKAQEAFLAQLQKATKOEILTYYNK 120
QY 121 VYMSNGNGMGTAAQNYGKDLNNLSLPQALLAGMPQAPNOYDPSHPEAODRRNLVL 180
DB 121 VYMSNGNGMGTAAQNYGKDLNNLSLPQALLAGMPQAPNOYDPSHPEAODRRNLVL 180
QY 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNYLKEVINQVEEETGYNLT 240
DB 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNYLKEVINQVEEETGYNLT 240
QY 241 TGMDEVYTNVDEAOKHLMIDYNTDEVYVAPDELOVASTIVDSNGKVIQALGARHSSN 300
DB 241 TGMDEVYTNVDEAOKHLMIDYNTDEVYVAPDELOVASTIVDSNGKVIQALGARHSSN 300
QY 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTNPVYNNDR 360
DB 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTNPVYNNDR 360
QY 361 GFNGNITLQYALQOSRNPVAVETLNKVGILNRAKTFNLGLIDYPSIHVSNAISSNTTESD 420
DB 361 GFNGNITLQYALQOSRNPVAVETLNKVGILNRAKTFNLGLIDYPSIHVSNAISSNTTESD 420
QY 421 KRYGASSEKMAAAYAFANGGTYYPMYIHKVFSDGSEKESNVGTRAMKETTYAYMMTD 480

|||||
Db 421 KKGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYMTD 480
QY 481 MKTIVLYTGGRNAYLAWLPQAGKTGTSNTDEIENHIKTSQFVAPDELFACTYRKYSM 540
Db 481 MKTIVLYTGGRNAYLAWLPQAGKTGTSNTDEIENHIKTSQFVAPDELFACTYRKYSM 540
QY 541 AVWTGYSNRLTPLYGNGLYVAAKAVYRSMYTLSEGSNPEDMNIPEGLYRNGEVEFRNGAR 600
Db 541 AVWTGYSNRLTPLYGNGLYVAAKAVYRSMYTLSEGSNPEDMNIPEGLYRNGEVEFRNGAR 600
QY 601 STWNSPAPQPPSTESSSSSDSSTSSSTSTPSTNNSTTTNNNTQOQNTTPTDOQNON 660
Db 601 STWNSPAPQPPSTESSSSSDSSTSSSTSTPSTNNSTTTNNNTQOQNTTPTDOQNON 660
QY 661 POPAP 666
Db 661 POPAP 666

RESULT 3

US-09-765-272-2
; Sequence 2, Application US/09765272
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8512
; TELEFAX: (301) 309-8504
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-272-2

Query Match 100.0%; Score 3484; DB 21; Length 666;
Best Local Similarity 100.0%; Pred. No. 3e-289;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADGSRERRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRILGAFLRN 60
Db 1 KIYDNKNOLIADGSRERRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRILGAFLRN 60
QY 61 LOSNSLOGGSTLTQOLIKILYFSTSDQTSRKAEAWLAIQLEKATQOEILTYIYINK 120

|||||
Db 61 LOSNSLOGGSTLTQOLIKILYFSTSDQTSRKAEAWLAIQLEKATQOEILTYIYINK 120
QY 121 VVMSNMCYMGTAQNAQNYVYGDNLNLSPOLALLAGPQAPNVDPSHPAADDRLNLY 180
Db 121 VVMSNMCYMGTAQNAQNYVYGDNLNLSPOLALLAGPQAPNVDPSHPAADDRLNLY 180
QY 181 SEKNNGYISAEQYKAVNPTPIDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240
Db 181 SEKNNGYISAEQYKAVNPTPIDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240
QY 241 TGMDEVYTNVDQEAQKHLMDIYNTDEXVAYPPDELAQVASTIVDSNGKVIAQLGARRHSSN 300
Db 241 TGMDEVYTNVDQEAQKHLMDIYNTDEXVAYPPDELAQVASTIVDSNGKVIAQLGARRHSSN 300
QY 301 VSFGINQAVETNNDGKSTKPTIDYAPALEGYDYSTATIVHDEPNYPTNTPPVNMDR 360
Db 301 VSFGINQAVETNNDGKSTKPTIDYAPALEGYDYSTATIVHDEPNYPTNTPPVNMDR 360
QY 361 GYFGNTITLOYALQOSRNVPAVETLNKVGILNRAKTFPLNGIDYPSIHYSAISSNTTESD 420
Db 361 GYFGNTITLOYALQOSRNVPAVETLNKVGILNRAKTFPLNGIDYPSIHYSAISSNTTESD 420
QY 421 KKGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYMTD 480
Db 421 KKGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYMTD 480
QY 481 MKTIVLYTGGRNAYLAWLPQAGKTGTSNTDEIENHIKTSQFVAPDELFACTYRKYSM 540
Db 481 MKTIVLYTGGRNAYLAWLPQAGKTGTSNTDEIENHIKTSQFVAPDELFACTYRKYSM 540
QY 541 AVWTGYSNRLTPLYGNGLYVAAKAVYRSMYTLSEGSNPEDMNIPEGLYRNGEVEFRNGAR 600
Db 541 AVWTGYSNRLTPLYGNGLYVAAKAVYRSMYTLSEGSNPEDMNIPEGLYRNGEVEFRNGAR 600
QY 601 STWNSPAPQPPSTESSSSSDSSTSSSTSTPSTNNSTTTNNNTQOQNTTPTDOQNON 660
Db 601 STWNSPAPQPPSTESSSSSDSSTSSSTSTPSTNNSTTTNNNTQOQNTTPTDOQNON 660
QY 661 POPAP 666
Db 661 POPAP 666

RESULT 4

US-09-583-110-3973
; Sequence 3973, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATN00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3973
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3973

Query Match 100.0%; Score 3483; DB 19; Length 719;
Best Local Similarity 99.8%; Pred. No. 4.2e-289;
Matches 665; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADGSRERRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRILGAFLRN 60

|||||
Db 54 KLYDKNKQNLIALDGSERRVNAQANDIPDLVKAIVSIEDHREFDHGIDITIRILGAFLRN 113
Qy 61 LOSNSLOGGSTLTQOLIKLITYFSTSDQTSRKQAEMLAIQLEOKATKOEILTYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLITYFSTSDQTSRKQAEMLAIQLEOKATKOEILTYINK 173
Qy 121 VYMSNGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPAQAODRRNLVL 180
Db 174 VYMSNGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPAQAODRRNLVL 233
Qy 181 SEMKNQGYISAQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLLT 240
Db 234 SEMKNQGYISAQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLLT 293
Qy 241 TGMADYTVNDQEAOKHMDIYNTDEVYAPDDELQVASTIVDVSNGKYIAOLGARHOSN 300
Db 294 TGMADYTVNDQEAOKHMDIYNTDEVYAPDDELQVASTIVDVSNGKYIAOLGARHOSN 353
Qy 301 VSEGINQAVETNRDMSGTMRPTTDYAPALEYGVYDSTATIVHDEPYNPGTTPYVNMMDR 360
Db 354 VSEGINQAVETNRDMSGTMRPTTDYAPALEYGVYDSTATIVHDEPYNPGTTPYVNMMDR 413
Qy 361 GYFGNITLQYALQOSRNPAYETLNKVGILNRKATPLNGLGIDYPSIHSNAISSNTESD 420
Db 414 GYFGNITLQYALQOSRNPAYETLNKVGILNRKATPLNGLGIDYPSIHSNAISSNTESD 473
Qy 421 KKYGASSEKMAAAYAAAFANGGTYKPMYTHKVVFSDSGEKESNVGTRAMKETTAVMMD 480
Db 474 KKYGASSEKMAAAYAAAFANGGTYKPMYTHKVVFSDSGEKESNVGTRAMKETTAVMMD 533
Qy 481 MKKTVLYTGTGNAYLAWLPQAGKGTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTVLYTGTGNAYLAWLPQAGKGTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRKYSM 593
Qy 541 AVMTGYSNRLTPLYNGGLTVAAKYVRSMTYLSSEGSNPEDMNIPGGLYRNGEFVKNGAR 600
Db 594 AVMTGYSNRLTPLYNGGLTVAAKYVRSMTYLSSEGSNPEDMNIPGGLYRNGEFVKNGAR 653
Qy 601 STWNSPAPQPPSTESSSSSDSSTSSSTPTNNSTTTNNNTQOOSNTTTPDOONON 660
Db 654 STWNSPAPQPPSTESSSSSDSSTSSSTPTNNSTTTNNNTQOOSNTTTPDOONON 713
Qy 661 PQPAP 666
Db 714 PQPAP 719

RESULT 5
US-09-107-433-3705
: Sequence 3705, Application US/09107433
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NOCLEIC ACID AND AMINO ACID
: THERAPEUTICS
: NUMBER OF SEQUENCES: 5206
: CORRESPONDENCE ADDRESS:
: ADDRESS: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER: <Unknown>
: OPERATING SYSTEM: <Unknown>
: SOFTWARE: <Unknown>
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107.433
: FILING DATE: 30-Jun-1998
: PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arianello, Pamela Deneke
REGISTRATION NUMBER: 40, 489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3705:
SEQUENCE CHARACTERISTICS:
LENGTH: 721 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...721
SEQUENCE DESCRIPTION: SEQ ID NO: 3705:
US-09-107-433-3705

Query Match 100.0%; Score 3483; DB 15; Length 721;
Best Local Similarity 99.8%; Pred. No. 4.3e-289;
Matches 665; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KLYDKNKQNLIALDGSERRVNAQANDIPDLVKAIVSIEDHREFDHGIDITIRILGAFLRN 60
Db 56 KLYDKNKQNLIALDGSERRVNAQANDIPDLVKAIVSIEDHREFDHGIDITIRILGAFLRN 115
Qy 61 LOSNSLOGGSTLTQOLIKLITYFSTSDQTSRKQAEMLAIQLEOKATKOEILTYINK 120
Db 116 LOSNSLOGGSTLTQOLIKLITYFSTSDQTSRKQAEMLAIQLEOKATKOEILTYINK 175
Qy 121 VYMSNGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPAQAODRRNLVL 180
Db 176 VYMSNGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPAQAODRRNLVL 235
Qy 181 SEMKNQGYISAQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLLT 240
Db 236 SEMKNQGYISAQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLLT 295
Qy 241 TGMADYTVNDQEAOKHMDIYNTDEVYAPDDELQVASTIVDVSNGKYIAOLGARHOSN 300
Db 296 TGMADYTVNDQEAOKHMDIYNTDEVYAPDDELQVASTIVDVSNGKYIAOLGARHOSN 355
Qy 301 VSEGINQAVETNRDMSGTMRPTTDYAPALEYGVYDSTATIVHDEPYNPGTTPYVNMMDR 360
Db 356 VSEGINQAVETNRDMSGTMRPTTDYAPALEYGVYDSTATIVHDEPYNPGTTPYVNMMDR 415
Qy 361 GYFGNITLQYALQOSRNPAYETLNKVGILNRKATPLNGLGIDYPSIHSNAISSNTESD 420
Db 416 GYFGNITLQYALQOSRNPAYETLNKVGILNRKATPLNGLGIDYPSIHSNAISSNTESD 475
Qy 421 KKYGASSEKMAAAYAAAFANGGTYKPMYTHKVVFSDSGEKESNVGTRAMKETTAVMMD 480
Db 476 KKYGASSEKMAAAYAAAFANGGTYKPMYTHKVVFSDSGEKESNVGTRAMKETTAVMMD 535
Qy 481 MKKTVLYTGTGNAYLAWLPQAGKGTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRKYSM 540
Db 536 MKKTVLYTGTGNAYLAWLPQAGKGTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRKYSM 595
Qy 541 AVMTGYSNRLTPLYNGGLTVAAKYVRSMTYLSSEGSNPEDMNIPGGLYRNGEFVKNGAR 600
Db 596 AVMTGYSNRLTPLYNGGLTVAAKYVRSMTYLSSEGSNPEDMNIPGGLYRNGEFVKNGAR 655
Qy 601 STWNSPAPQPPSTESSSSSDSSTSSSTPTNNSTTTNNNTQOOSNTTTPDOONON 660

Db 656 STWNSPAPQPPSTESSSSSDSSTSSSTPSTNNSTTTNNNTTQOSNTTPOQNON 715
Qy 661 POPAP 666
Db 716 POPAP 721

RESULT 6

PCT-US02-03987-13423
; Sequence 13423, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits
; FILE REFERENCE: ELITTRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US02-03987-13423

Query Match 99.5%; Score 3466; DB 1; Length 719;
Best Local Similarity 99.4%; Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDNKNOLIADGSERRVNAQANDIPTDLVKAIVSIEDHRFPDHRGIDTIRILGAFARN 60
Db 54 KIYDNKNOLIADGSERRVNAQANDIPTDLVKAIVSIEDHRFPDHRGIDTIRILGAFARN 113
Qy 61 LOSNSLOGSGSTLTQOLIKLTYFSTSDQTSRKAQEWALAIQLEKATQOELITYINK 120
Db 114 LOSNSLOGSGSALTQOLIKLTYFSTSDQTSRKAQEWALAIQLEKATQOELITYINK 173
Qy 121 VYMSNGNYGQTAQNYVYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVLT 180
Db 174 VYMSNGNYGQTAQNYVYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVLT 233
Qy 181 SEKKNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEVINOVEEFTGYNLLT 240
Db 234 SEKKNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEVINOVEEFTGYNLLT 293
Qy 241 TGMADVNTNDOEQAKHLMIDYNTDEVYAYPDDDELQVASTIVDSNGKVIAQIGARRQSSN 300
Db 294 TGMADVNTNDOEQAKHLMIDYNTDEVYAYPDDDELQVASTIVDSNGKVIAQIGARRQSSN 353
Qy 301 VSEGINQAVETNRDMSYMKPTIDYAPALEYGYVSTATIVHDEPYNYPGTNTPVYNMWR 360
Db 354 VSEGINQAVETNRDMSYMKPTIDYAPALEYGYVSTATIVHDEPYNYPGTNTPVYNMWR 413
Qy 361 GYFNGNTLQYALQOOSRNVAVENTLKVGNRAKTFPLNGGIDVPSIHYSAISSNTTESD 420
Db 414 GYFNGNTLQYALQOOSRNVAVENTLKVGNRAKTFPLNGGIDVPSIHYSAISSNTTESD 473
Qy 421 KRYGASSEKMAAAYAFANGGYTYKPMYIHKVYFSDGSEKEFSNVGTTRAKKETTAAMMTD 480
Db 474 KRYGASSEKMAAAYAFANGGYTYKPMYIHKVYFSDGSEKEFSNVGTTRAKKETTAAMMTD 533
Qy 481 MKKTYVLTGTGRNAVLAAMLPOAGKTGTSNTTDEIENHIKTSQFVAPDELFACTRYKYSM 540
Db 534 MKKTYVLTGTGRNAVLAAMLPOAGKTGTSNTTDEIENHIKTSQFVAPDELFACTRYKYSM 593
Qy 541 AVMTGSNLTPLVNGGLVAAKVVYSMMTYLSEGSNPEEDMNTPEGLYRNGEEVFENGAR 600
Db 594 AVMTGSNLTPLVNGGLVAAKVVYSMMTYLSEGSNPEEDMNTPEGLYRNGEEVFENGAR 653
Qy 601 STWNSPAPQPPSTESSSSSDSSTSSSTPSTNNSTTTNNNTTQOSNTTPOQNON 660

Db 654 STWSSPAPQPPSTESSSSSDSSTSSSTPSTNNSTTTNNNTTQOSNTTPOQNON 713
Qy 661 POPAP 666
Db 714 POPAP 719

RESULT 7

US-09-815-242-13423
; Sequence 13423, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423

Query Match 99.5%; Score 3466; DB 22; Length 719;
Best Local Similarity 99.4%; Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDNKNOLIADGSERRVNAQANDIPTDLVKAIVSIEDHRFPDHRGIDTIRILGAFARN 60
Db 54 KIYDNKNOLIADGSERRVNAQANDIPTDLVKAIVSIEDHRFPDHRGIDTIRILGAFARN 113
Qy 61 LOSNSLOGSGSTLTQOLIKLTYFSTSDQTSRKAQEWALAIQLEKATQOELITYINK 120
Db 114 LOSNSLOGSGSALTQOLIKLTYFSTSDQTSRKAQEWALAIQLEKATQOELITYINK 173
Qy 121 VYMSNGNYGQTAQNYVYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVLT 180
Db 174 VYMSNGNYGQTAQNYVYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVLT 233
Qy 181 SEKKNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEVINOVEEFTGYNLLT 240
Db 234 SEKKNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEVINOVEEFTGYNLLT 293
Qy 241 TGMADVNTNDOEQAKHLMIDYNTDEVYAYPDDDELQVASTIVDSNGKVIAQIGARRQSSN 300
Db 294 TGMADVNTNDOEQAKHLMIDYNTDEVYAYPDDDELQVASTIVDSNGKVIAQIGARRQSSN 353
Qy 301 VSEGINQAVETNRDMSYMKPTIDYAPALEYGYVSTATIVHDEPYNYPGTNTPVYNMWR 360

Db 354 VSEFGINQAVETNRDMSGSTMKPTTIDYAPALEYGVYESTATIVHDEPNYNGTNPVYNMNR 413
Qy 361 GYFGNTTQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNAISSNTESD 420
Db 414 GYFGNTTQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNAISSNTESD 473
Qy 421 KRYGASSEKMAAAYAFANGGTYRPMYTHKVVFSGDSSEKESNNGTRAMKETTAAYMTD 480
Db 474 KRYGASSEKMAAAYAFANGGTYRPMYTHKVVFSGDSSEKESNNGTRAMKETTAAYMTD 533
Qy 481 MKKTIVLTGTCGNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFGYTRKYSM 540
Db 534 MKKTIVLTGTCGNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFGYTRKYSM 593
Qy 541 AVMTGYSNRLEPLVGNGLTVAAYKVRSMWMTYSEGSNPEDMNIPEGLYRNGEYFRKNGAR 600
Db 594 AVMTGYSNRLEPLVGNGLTVAAYKVRSMWMTYSEGSNPEDMNIPEGLYRNGEYFRKNGAR 653
Qy 601 STWNSPAPQOPPSTESSSSSDSSTSSQSSSTPTNNSTTTNPNNNTQOANTTPOQONON 660
Db 654 STWNSPAPQOPPSTESSSSSDSSTSSQSSSTPTNNSTTTNPNNNTQOANTTPOQONON 713
Qy 661 POPAOP 666
Db 714 POPAOP 719

RESULT 8
US-10-072-851-13423
: Sequence 13423, Application US/10072851
: GENERAL INFORMATION:
: APPLICANT: Carr, Grant J.
: APPLICANT: Xu, H. Howard
: APPLICANT: Foulkes, J. Gordon
: APPLICANT: Zamudio, Carlos
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Twilick, John D.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Roemer, Terry
: APPLICANT: Jiang, Bo
: APPLICANT: Boone, Charles
: APPLICANT: Bussey, Howard
: TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
: FILE REFERENCE: ELITRA.028A
: CURRENT APPLICATION NUMBER: US/10/072,851
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: 60/267,636
: NUMBER OF SEQ ID NOS: 15811
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13423
: LENGTH: 719
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
: US-10-072-851-13423

Query Match 99.5%; Score 3466; DB 24; Length 719;
Best Local Similarity 99.4%; Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDKNQLIADLGSERVRNAQANDIPDLYKAVISIEDHREFDRGIDITRILGAFLENN 60
Db 54 KIYDKNQLIADLGSERVRNAQANDIPDLYKAVISIEDHREFDRGIDITRILGAFLENN 113
Qy 61 LOSNSLOGGSLTQOLIKLYTFSTISDOTISRKAQEWMLAIQLEOKATKOEILTYIYNK 120
Db 114 LOSNSLOGGSLTQOLIKLYTFSTISDOTISRKAQEWMLAIQLEOKATKOEILTYIYNK 173

Qy 121 VYMSNGNVMQTAQONVYKGLDNLNLSLPOLALLAGMPQAPNOYDPYSHPAADOORNLVL 180
Db 174 VYMSNGNVMQTAQONVYKGLDNLNLSLPOLALLAGMPQAPNOYDPYSHPAADOORNLVL 233
Qy 181 SEMKNQGYISABOYEKAVNPTTIDGLOSLSKASNPAYMDNLYKEVINQVEEBETGYNLT 240
Db 234 SEMKNQGYISABOYEKAVNPTTIDGLOSLSKASNPAYMDNLYKEVINQVEEBETGYNLT 293
Qy 241 TGMIDYITNDQAKHMDIYNTDEYVATPDELOVASTIVVSNCKVYLAQIGARQSSN 300
Db 294 TGMIDYITNDQAKHMDIYNTDEYVATPDELOVASTIVVSNCKVYLAQIGARQSSN 353
Qy 301 VSEFGINQAVETNRDMSGSTMKPTTIDYAPALEYGVYESTATIVHDEPNYNGTNPVYNMNR 360
Db 354 VSEFGINQAVETNRDMSGSTMKPTTIDYAPALEYGVYESTATIVHDEPNYNGTNPVYNMNR 413
Qy 361 GYFGNTTQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNAISSNTESD 420
Db 414 GYFGNTTQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNAISSNTESD 473
Qy 421 KRYGASSEKMAAAYAFANGGTYRPMYTHKVVFSGDSSEKESNNGTRAMKETTAAYMTD 480
Db 474 KRYGASSEKMAAAYAFANGGTYRPMYTHKVVFSGDSSEKESNNGTRAMKETTAAYMTD 533
Qy 481 MKKTIVLTGTCGNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFGYTRKYSM 540
Db 534 MKKTIVLTGTCGNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFGYTRKYSM 593
Qy 541 AVMTGYSNRLEPLVGNGLTVAAYKVRSMWMTYSEGSNPEDMNIPEGLYRNGEYFRKNGAR 600
Db 594 AVMTGYSNRLEPLVGNGLTVAAYKVRSMWMTYSEGSNPEDMNIPEGLYRNGEYFRKNGAR 653
Qy 601 STWNSPAPQOPPSTESSSSSDSSTSSQSSSTPTNNSTTTNPNNNTQOANTTPOQONON 660
Db 654 STWNSPAPQOPPSTESSSSSDSSTSSQSSSTPTNNSTTTNPNNNTQOANTTPOQONON 713
Qy 661 POPAOP 666
Db 714 POPAOP 719

RESULT 9
US-09-107-532-5667
: Sequence 5667, Application US/09107532
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO
: TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
: TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
: NUMBER OF SEQUENCES: 7308
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER:
: OPERATING SYSTEM:
: SOFTWARE:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107,532
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/085598
: FILING DATE: May 14, 1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/051571
: FILING DATE: July 2, 1997
: ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5667:
SEQUENCE CHARACTERISTICS:
LENGTH: 823 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...823
US-09-107-532-5667

Query Match 53.5%; Score 1862.5; DB 15; Length 823;
Best Local Similarity 53.9%; Pred. No. 7e-150;
Matches 371; Conservative 109; Mismatches 179; Indels 29; Gaps 8;

1 KYIDKNOIADI GSRVNAQANDIPTDLVKAIVSIEDHREFDRHGIDITRILGAFLN 60
115 KLYTODGELEFEDGAEKREKISANELPKTLEDIAIVSEDRFRFKHGVDPRIIGSALS 174
61 LOSNSLOGSSTLTQOLIKLTYFSTSDOTISRKAEAMLAIOLEOKATQOELITYYINK 120
175 FTSGGILOGSSTLTQOLIKLTSFSTSAEDOTLKRKAQAMMAVRLBOKKSKOELITYYINK 234
121 VYMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRLVY 180
235 VYMSNGLYGMEIASSEMYFEKLSLSELPOTALLAGMPQAPSAVDYVYPDOAKKRDTVL 294
181 SEMKNQGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVYINOVEETGNLTL 240
295 YTMLONEKISQREYDQAVNVPYTDGLQELTQSDDNTKIYDNYKVEYINVEQETDKNVT 354
241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAVPDEDLQVASTIVDSNGKVIAGLGAHQSSN 300
355 DLEIYTNLDLDAOKKLYIVMTDOYVSPDEDMQVASTLIDTNGKVAQIGRIAD 414
301 VSEGINQAVETNRDMGSTMKPTIDYAPALEYGVYDSTATIVHDEPNYFGTNTPVYNNDR 360
415 YTLGNLNAVNTSRDFGSTRKPVTDYGAPEYLYK-STGKTIDAPYNEGSTSPVGNMND 473
361 GFNGNTTLOALQOOSRNVAVELINKVGLNRAKTFPLNGIGIYPSHYNSNATSSNTEED 420
474 QYMGTTTLQALYLSRNNVAVKLFNEVGSQVASFKNLNGISTHGNATSSNTEED 533
421 -KRYGASSSEKMAAAVAFANGGTYYKPMYIHKVFESEKSEFSNNGTAMKETTAYMMT 479
534 GKRYGASSLAKMAAAAFANGGTYYKPYVKNKIYFODGTEERYEPBGKRAMSETYMYT 593
480 DMKATVLYGTRNAVLAWLPOAGKTGTSNYDEETENHIKTSQVAPDELPAGYTRKYS 539
594 DILKTLITEGTGTNAQIAGLYOAGKTGTSNYDDEYAK-LGISSGVYPILEFAGYTPNYS 652
540 MAVVTGYSRLLPIVINGNLTVAAKYRSMMTYLSSEGSNEDMNNIPGLR-NOEFYKRN- 597
653 ISVWGTGYNKKMPTVSESHVASDVRRELMQVAVANVTMDMPGSLRVGGELYKDO 712
598 -GARSTWNSPAPQOP-----PSTESSSSSDSTSSQSSSTPTSTNNST----- 639
713 YIARSAIATPSTTIPSSSVVQTPGSGSTTETTTQSSSSTQSSSTASSSTTAETSEPA 772
640 --TTINNNNTQOQNT---TPDOQONP 661
773 SSTIVPSSSSSESSSTPSSSAPASSSEP 800

RESULT 10
US-09-107-532A-5667
Sequence 5667, Application US/09107532A
GENERAL INFORMATION:
APPLICANT: Lynn A Doucelte-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5667:
SEQUENCE CHARACTERISTICS:
LENGTH: 823 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...823
SEQUENCE DESCRIPTION: SEQ ID NO: 5667;
US-09-107-532A-5667

Query Match 53.5%; Score 1862.5; DB 15; Length 823;
Best Local Similarity 53.9%; Pred. No. 7e-150;
Matches 371; Conservative 109; Mismatches 179; Indels 29; Gaps 8;

1 KYIDKNOIADI GSRVNAQANDIPTDLVKAIVSIEDHREFDRHGIDITRILGAFLN 60
115 KLYTODGELEFEDGAEKREKISANELPKTLEDIAIVSEDRFRFKHGVDPRIIGSALS 174
61 LOSNSLOGSSTLTQOLIKLTYFSTSDOTISRKAEAMLAIOLEOKATQOELITYYINK 120
175 FTSGGILOGSSTLTQOLIKLTSFSTSAEDOTLKRKAQAMMAVRLBOKKSKOELITYYINK 234
121 VYMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRLVY 180
235 VYMSNGLYGMEIASSEMYFEKLSLSELPOTALLAGMPQAPSAVDYVYPDOAKKRDTVL 294
181 SEMKNQGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVYINOVEETGNLTL 240
295 YTMLONEKISQREYDQAVNVPYTDGLQELTQSDDNTKIYDNYKVEYINVEQETDKNVT 354
241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAVPDEDLQVASTIVDSNGKVIAGLGAHQSSN 300

```

Db      355 DGLEIYVTLMDLDAQKLIYDINTDQVYSYPDDEMQVASTLIDTNGKVKWQIGRHTAED 414A
QY      301 VSEFGINAVLEFNRMQGSTMKPIYDAPALEYGVYDSTATIYHDEPYNTPGCTNPPYVNMWR 360B
Db      415 VTLGNNILAVNTPSROFGSTPMKEPVTIDYGAPEFLYK -STGKTIIDAPNYEGSTPFGWMDN 473C
QY      361 GYFNGITLOVLAQOSRNPVAVETLNKGLNMAKFLUNGIGLDIADPYSIHYSNAISNTPTESD 420D
Db      474 QYMGITITRQVLALYSRNPVPAKFLNEBGSDKVASFELNGLIGESTIHQSNASISNTEEO 533E
QY      421 -KRYGASSEKMAAAVAAAFANGCTYKKPMYIHKVYFSDGSEKEFSGNVSTRAKMEKTAYMMT 479F
Db      534 GTKKGASLLKMAAAVAAAFANGCTYKKQYVANKIYFQGTETETYPDPDKTAMSETAAMIT 553G
QY      480 DMMKTVLITYGTGRNAYLAAMLPOAQKGTGSNTYDEIEENHIKTSQFVAPDELFAGYTERKYS 539H
Db      594 DILKDTIEGCTGNAQIAGLVQAKGTGSNTYDDEYAK -LGISGCVYDILFAGYTPNYS 652I
QY      540 MAVMTGYSNRLTPLYGVNGLYAAKVYRSMMTYLSGSGNPEPDWNIPEGLYR-NGEYVFKN- 597J
Db      653 ISVWTGYKKMTPTVTSSESHVASVYRELQVYASANTNDMDWEMPGLLIRVGCGLYKQD 712K
QY      598 -GARSTWMSRPAPOQ-----PTESSSSSSSDSTSOSSSTTTPBNTNST----- 639L
Db      713 YTARSNALITPTSTTIPSSSYVOTPGSSYTTETITTOSSSSTSOSESYAESSEKSETAETSEPA 772M
QY      640 ---TTPNNNTQOASNT---PDPQONOP 661N
Db      773 SSTTVPSSSSSEESTPSSSAPAPASSSEP 800O

```

```

RESULT 11
PCT-US02-03987-10728
: Sequence 10728, Application PC/WTUS0203987
: GENERAL INFORMATION:
: APPLICANT: Elitra Pharmaceuticals, Inc.
: TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
: TITLE OF INVENTION: Proliferation
: FILE REFERENCE: ELITRA.028VPC
: CURRENT APPLICATION NUMBER: PCT/US02/03987
: CURRENT FILING DATE: 2002-02-02
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: NUMBER OF SEQ ID NOS: 15811
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 10728
: LENGTH: 778
: TYPE: PRT
: ORGANISM: Enterococcus faecalis
PCT-US02-03987-10728

```

Query Match	50.3%	Score 1751;	DB 1;	Length 778;
Best Local Similarity	51.4%;	Pred. No. 2.4e-140;		
Matches 347;	Conservative 125;	Mismatches 193;	Indels 10;	Gaps 3.

[illegible]

Qy	241	TGMVYIVWDEADKHLMDIYNTDEYAAVPDDELQAVSTIVDSNGKVIYIQLQGLARHSSN	300
Db	321	DGLDIYTLMDMNAQKQLDYIDNSQOYAFPDKKQVASTIVDAISGVYRQIGRHLIPDD	360
Qy	301	VSEGINOAVETNRDMGSTMKPFTDYAPALEYGVYDSATYVHDEPYVPGCTNTPYVWMDR	360
Db	381	VQLGNNLAVNTRQVGVSTVKEIPMIGYALENLNT-SIGRLVWVKPRYKPTGTDIDVFNSDL	439
Qy	361	GYFGNITLOVALOOSRNVPAVETLNKYGULNARKFLNGLGIDPYSIHYSNAISGNTESD	420
Db	440	TYQGYITTKRAIMSRNTTAVQGTDEYKEMIFPIGLGIDYKNLNASNAISSNTSDVD	499
Qy	421	KKYGASSEKMAAAYAAFPANGTVYKPMYIHKVYFSDGSEKESNVGTRAKETTYAMT	479
Db	500	GDKYGISLKLAAAYAAFPANGVIYKRYVYKVFNDGTSVDYQDPDRAKRMKSTAYMT	559
Qy	480	DMKKTLYTGGRNAYLAWLPQAKCTGTSNNTDEIENHIKTSQFVAPDELFGYTRPKS	539
Db	560	DMLDVLNGSGFGFAGLIPGLLOAKKTGTSNNTDIDLARMCTTEGCIAPDSTFVGYYTHYA	619
Qy	540	MAVWTGYSNRLTPLYVNGNLVFAAVKYVRSMYTYLSEGSNPEDMNIPEGLYRNGEVEFNGA	599
Db	620	VSVMTGYNDRMTPYIQEYFYGLASIVYELMISYLSQNNSDWQVDPDSVAVGMDLYKDA	679
Qy	600	RSTWN-----SPAQQPSTSESSSSSDSSTSQSSSTTPSTNNSTTNPNNNTOQSN	651
Db	680	YEYVNVQVLPSTSSAQPOPESSSTVEESSTKEAESSSSSSESAPSSSEAPRSTEGOPASS	739
Qy	662	TTTPOQNONPOPAOP	666
Db	740	SSADOPATSEOPPEP	754

```

RESULT 12
US-09-815-242-10728
: Sequence 10728, Application US/095815242
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELTRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10728
: LENGTH: 778
: TYPE: PRT
: ORGANISM: Enterococcus faecalis
: US-09-815-242-10728

```

ORGANISM: Enterococcus faecalis
US-09-815-242-10728

Query Match	50.3%;	Score 1751;	DB 22;	Length 778;
Best Local Similarity	51.4%;	Pred. No. 2.4e-140;		
Matches 347;	Conservative 125;	Mismatches 193;	Indels 10;	Gaps 3

Qy	1	KIYNKNOIJDJLSSERRVANAQADIPDIDLVAKYA:JEDHFRPHRDIIDTRIIGALRN	60
Db	81	KLYIINNEIFEDDLAEKRELIQPMDVQOLKDAIVSDEDRFKYHIQYDPIRIIGASLSN	140
Qy	61	LOSNSLOGGSTALLTOOLIKLYFSTSDQOTISRAQOEAMLAIOLEOKATKOEILTYIYNK	120
Db	141	VKNNGLOGGSTALLTOOLIKLYFSTESQDQFLKRAQOEAAMVAVRLEREKSKEEILTYIYNK	200
Qy	121	VYNSNGNGMOTAAONNYGYGDNLNLSI:POLALLAGMGAOPNDQVYRSHPEAODRRULY	180
Db	201	VYMANGEFGMTAAENYGYKHLSELDPQRLXLAGMGAOPASTDYPYKPTPAERDRUVL	260
Qy	181	SEMNQOYISAEQYERKAVNTPITDTGLOLSKASASNPAYMMNLYKEVINQOEETGYALLT	240
Db	261	YTYMDNNKISAEYERKAKATPIDEGLVPLKASDNRKVDNRYKVEVINEYKAKTGKNVYT	320
Qy	241	TGMDVYTNVQDEAKHLMIDYINDEYEAAYIPDDELQVASTIVDSNGKVIYIQLGARHOSSN	300
Db	321	DGLDIYTNLDMNNAKOLXYDIAVNSQYVAFPPDKKQVASTIVDASGGVRAOIGRH1PPD	380
Qy	301	VSEFINOAVETNRPMGSTMKPITPYAPALLEGYVDSATVYHDEPNYVPGTNPVYMMDR	360
Db	381	VOLGNMLAVNTRQVGSYVAKPIMDYGAILELNAN-SIGRRLMVDKPYTTPCTGIDIVFNSDL	439
Qy	361	GYEENITLOVALQOOSRNVPAVELTNKYGLNBAKCTFLNGLGIDPYSIHYNSAISNTESD	420
Db	440	TYQGVITMRRAIMGSRMTTAAQTFDEYEGKENIMFIFGLGIDYKNLFEASNAISNTESD	499
Qy	421	-KRYGASEKKAAYAAFPANGCTYKPMYIHKYVESFQSGSEKESNVGSTRAMKEPTAYMMT	479
Db	500	GDKTIGISSKLIAAAYAAFPANNGLYINAKPYVAKYVFENDGTSYDIOPPDKRAMKSDTAYMMT	558
Qy	480	DMKATVLYTGTGRNAYLAMLPOACRKTGTSNVTDEIEINHIKTSQFVAPDELFAGYTRKYS	539
Db	560	DMKDYLVNGGTGFGNGLPGLIOAAKTKGNSNTDEDLARMGTTEKGIAIDPSTFVGYTHYA	619
Qy	540	MAYVTGYSNRLTPLVGNGLYVAAVYYSMMTYLSEGSNPEDMNIPEDLIRNGEYVFRNGA	599
Db	620	VSVMTGYNDRTPTIYOEYEGIASVYREIMSYLSQANSNDWQOPDSVYVRGNELIYKDA	679
Qy	600	RSTNN-----SPAQOPSTRESSSSSSSSSTSSSSSTPTSTNNSTTNNPNNNTOOSN	651
Db	680	YEVYNNQVLPSTTSSAQOPRESSSTVESSSTKEALSSSSSSSEAPASSSEAPSTTEOPASS	739
Qy	652	TTPOOQONPOPAOP	666
Db	740	SSAEOPATSEOPPEP	754

RESULT 13
 US-10-072-851-10728
 : Sequence 10728, Application US/10072851
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Carr, Grant J.
 : APPLICANT: Xu, H. Howard
 : APPLICANT: Foulkes, J. Gordon
 : APPLICANT: Zamudio, Carlos
 : APPLICANT: Haselbeck, Robert
 : APPLICANT: Ohlsen, Karl L.
 : APPLICANT: Zyskind, Judith W.
 : APPLICANT: Wall, Daniel
 : APPLICANT: Trawick, John D.
 : APPLICANT: Yamamoto, Robert T.
 : APPLICANT: Roemer, Terry
 : APPLICANT: Jiang, Bo
 : APPLICANT: Boone, Charles
 : APPLICANT: Bussey, Howard
 :
 : TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
 :
 : TITLE OF INVENTION: Proliferation

```

; FILE REFERENCE: ELITRA 028A
; CURRENT APPLICATION NUMBER: US/10/072.851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10728
;
; LENGTH: 778
;
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-10-072-851-10728

```

Query Match	50.3%	Score 1751	DB 24,	Length 778;
Best Local Similarity	51.4%;	Pred. No. 2.4e-140;		
Matches 347; Conservative	125;	Mismatches 193;		
				Indels 10; Gaps 3

Oy	1	K IYDKRNKLLIDLTSESRVYNAQANDIPDVLAKAVSTJEDHRFPQHRIDPTRLIGALRN	60
Db	81	KLYDNNHIFEDLGAERKELLQRPDVPOLLKDALVSYEDRFRFKHIGVDPTRLIGALSIN	140
Oy	61	LOSNSLOGGSTLTOOLIKLYEFSTSTSDQTSRKADFAMLAIDLEQAKTMOELITYINK	120
Db	141	VKNGLGLOGSTLTOOLIKLSTFYSTKESDQTLKRAQCEAMAVBLREKSEKEELITYINK	200
Oy	121	YVMSNGNCGMOTANONYGRDLNNLSL.POLALLAGMOPANOVDPSYSHPEAADRRRLYL	180
Db	201	YVMAGGFYGMETAEENYKGLSELDPQRLALGMPADNSIDPYTKPPTAERRDVYL	260
Oy	181	SEMKNOGYISAEOYERKAVNPFITDLOSLKSSASNYPRPYMNYLKEVINOYEETGYNLLT	240
Db	261	YTMVDNKKISAEYERKAKATPIDEGVLPAKASDNRRKVDNYNVEYINEKAKTKRWYT	320
Oy	241	TGMVYTTWVDEDAOKHLMIDYINTEBEYAYIPDDELOVASTIVDYSNGKVIYQOLGARQSSN	300
Db	321	DGLDLYTNLDMNNAOKOLYDINSQYVAFPPDKMQVASTYIDVASCQVRAQICGRHLPDS	380
Oy	301	VSFGINOAVETNRPMGSTMKPITTYAPALALEGYVDSTATYVHDEPPYVPGTNPVYWMDR	360
Db	381	VOLGNMLAVNTRQDVGSTVYKRIIMPYGAILEMLN - STGRLLMDKPKTYPGTIDIDVFNSDL	439
Oy	361	GYFGNITLOYALQOOSRNVPAVETLNKYGKLNRAKTFELNGCIDPYSIHYSNAISSNTESD	420
Db	440	TYOGYTERRRAIMSRMNTTAAQUTDEDEGKEINIMFICLGDYKNLEASNAISSNTSDVD	499
Oy	421	KKYGASSEKMAAAYAAPFANGCTYKRYMYTHKYVFSQSKESKESPNVGTTRAMKETTVMMT	479
Db	500	GDKTKGISLKLAAAYAAPANNGITINPKPYVAKVYFNGTSDYDIDOPDKRAMKOSTALMMT	559
Oy	480	DMKTKVLYYTGGRNAYLAMLPOAGKRTGTSNYTDEIEINHIKTSOFVAPDELFGYTRAKYS	539
Db	560	DMLKDVILNGGTFGNCAIPGLIOAKTGTSNYTDEDLARMGCTTEKGIAIPDSTFGYGTTHVA	619
Oy	540	MAVWVTGYSNRLTPLVGNGLIYAAAVYYSMMTYLSEGSNPEDMNPBELYINGEFVFRNGA	599
Db	620	VSWMTGYMDRMTPLYQEYEGYGLASOYVEHIMSYLSQANSNDMWQPDSDVYVWGNELYYKDA	679
Oy	600	RSTWN-----SPAQOPSTSESSSSSDSPSTSGOESTTPTSTNNSTTTPNNNTQOOSN	651
Db	680	YEVPNVOYLPSTSSAPQPESSSYVEBSSTREADSSSSSSSEASAPSSSEAPSTBQPASS	739
Oy	652	TTTPOQONOPQAPOP	666
Db	740	SSAEQOPATSEQPEP	754

```

RESULT 14
US-09-134-000-4939
; Sequence 4939, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS

```

```
; TITLE OF INVENTION: FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-005
; CURRENT APPLICATION NUMBER: US/09/134.000A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 6810
; SEQ ID NO 4939
; LENGTH: 789
; TYPE: PR1
; ORGANISM: Enterococcus faecalis
US-09-134-000-4939
```

```
Query Match          50.3%; Score 1751; DB 15; Length 789;
Best Local Similarity 51.4%; Pred. No. 2.4e-140;
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;
```

```
QY 1 KIYDNKNOIADLIGSERVNAQANDIPTDLKAIYSIEDHFRFDRGIDTRIIGAFLRN 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 92 KLYDINNIEFDLGEKRELQPNQVLPOLLDALVSEDRRKYKHIGVDPRIITISALSN 151
QY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQITSRKAQEMALAIQLEOKATKQEIILTYINK 120
   :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 152 VKNGLOGSSTLTQOLIKLSTFKESQDITLKRQAEMAMAVRLERESKEIILTYINK 211
QY 121 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPAPNQDYPYSHPEAAQDRRLVL 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 212 VYMANGYMETAAENYKHLSELDELPTQALLAGMPAPNSYDPTAKERDQVYL 271
QY 181 SEMKNOGIISAEOYEKANPTITDGLQSLKASNPAYMDNYLKEVINQVEETGYNLLT 240
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 272 YTMVDNKKISAKEYEKAATPIDGLVPLKASDNRKVVADYVKEVINEVAKTKGNVYT 331
QY 241 TGMVYTNVDDQAKHLMIDYNTDEYVAYPPDELQVASTIVDVNSGKYIAQLGARHSSN 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 332 DGLDIYTNLDMNAQKQLYDIYNSQYVAFPPDKMQVASTIVDVNSGQVRAQIGRHIPDD 391
QY 301 VSEGINQAVETNRDMGSTMPTDYAPALEXYVDSTATIVHDEBYNPGTNPVYNDMR 360
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 392 VOLNNNLAVNTOBQVGVSTKPIMDYGAPEINLNY-STGLRLAWDKTKYKPGTDIDVFNSDL 450
QY 361 GYFENITLOVALQOSRNPAYETLNVKLNRAKTFNLGLGIDYPSIHSNMISSSTTSD 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 451 TYQGYITMRRAIMGSRNTTAQOTFDEVGKENIMPIKGLGIDYKLNLSNMISSSTTSD 510
QY 421 -KKYGASSEKMAAAYAAAFANGCTYKPMYIHKVYVSDSESEKSNVGRAKETAYMYT 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 511 GDXGIGISLKLAAAYAAAFANNGIYKPYKVFVNDGTSVDYQDGRKRAKDSATAYMYT 570
QY 480 DMKATVLTIGGRNAYLAMLQAGKTGTSNTDEIENHITSQFVAPDELPAGYTRKYS 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 571 DMLDVLNGGTGFNGALPGLIOAAKTGTSNTDELDLARMGTTEKGIADSTFVGYTTHYA 630
QY 540 MAVVTGYSNRLTPLYGNGLTVAARVYRSMNTYLSGSSNPEDMNTPEGLYRNGEYFNKGA 599
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 631 VSVNTGINDRNTPLYOEYTGIASDVRITKSYLSQNVSNDDWQVDSVYRVGNELLYVADA 690
QY 600 RSTWN-----SPAQPPSTESSSSSDSTSQSSSTPTSTNNSTPTNNNTQOOSN 651
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 691 YEYVNVQVLPSTTSAPQESSSTVESSSTKEASSSSSSSESAPSSSEAPRSTPDQPASS 750
QY 652 TTPDOQNONPOPAOP 666
   ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 751 SSAPQAPTSQPPPEP 765
```

```
RESULT 15
US-08-116-541-4
; Sequence 4, Application US/08116541
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Pearce, Barbara J.
; APPLICANT: Toumanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
```

```
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/116,541
; FILING DATE: 19930901
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 343-1684
; TELEFAX: 201 343-1684
; TELE: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-116-541-4
```

```
Query Match          46.5%; Score 1621; DB 5; Length 320;
Best Local Similarity 99.7%; Pred. No. 7.2e-130;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 KIYDNKNOIADLIGSERVNAQANDIPTDLKAIYSIEDHFRFDRGIDTRIIGAFLRN 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5 KIYDNKNOIADLIGSERVNAQANDIPTDLKAIYSIEDHFRFDRGIDTRIIGAFLRN 64
QY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQITSRKAQEMALAIQLEOKATKQEIILTYINK 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 LOSNSLOGSSTLTQOLIKLTYFSTSDQITSRKAQEMALAIQLEOKATKQEIILTYINK 124
QY 121 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPAPNQDYPYSHPEAAQDRRLVL 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 125 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPAPNQDYPYSHPEAAQDRRLVL 184
QY 181 SEMKNOGIISAEOYEKANPTITDGLQSLKASNPAYMDNYLKEVINQVEETGYNLLT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 185 SEMKNOGIISAEOYEKANPTITDGLQSLKASNPAYMDNYLKEVINQVEETGYNLLT 244
QY 241 TGMVYTNVDDQAKHLMIDYNTDEYVAYPPDELQVASTIVDVNSGKYIAQLGARHSSN 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 245 TGMVYTNVDDQAKHLMIDYNTDEYVAYPPDELQVASTIVDVNSGKYIAQLGARHSSN 304
QY 301 VSEGINQAVETNRDMG 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 305 VSEGINQAVETNRDMG 320
```

```
Search completed: August 11, 2002, 04:30:05
Job time: 801 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 04:23:44 ; Search time 142.66 Seconds

(Without alignments)
818.374 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484
Sequence: 1 KIYDKKNLIDLGSESRVYN.....TQGSMTTPDQGNQNPAP 666Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 528882 seqs, 175299045 residues

Total number of hits satisfying chosen parameters: 528882

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents AA New: *
1: /cgn2_6/ptodata/2/paa/PCOT_NEW_COMB.pep: *
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep: *
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep: *
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep: *
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep: *
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep: *
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	666	4 US-08-961-083-2	Sequence 2, Appli
2	529.5	15.2	821	5 US-09-769-744A-122	Sequence 122, App
3	526	15.1	789	5 US-09-540-209B-8941	Sequence 8941, Ap
4	384	11.0	790	5 US-09-540-209B-9358	Sequence 9358, Ap
5	162	4.6	875	5 US-09-914-543-6	Sequence 6, Appli
6	161	4.6	738	6 US-10-179-131-10183	Sequence 10183, A
7	160.5	4.6	2893	5 US-09-882-227-522	Sequence 522, App
8	159	4.6	615	5 US-09-882-327-408	Sequence 408, App
9	153.5	4.4	930	5 US-09-200-650D-3	Sequence 3, Appli
10	152	4.4	973	6 US-10-179-131-5286	Sequence 5286, Ap
11	148.5	4.3	2283	1 PCT-US02-19220-4	Sequence 4, Appli
12	148	4.2	188	5 US-09-540-209B-6592	Sequence 6592, Ap
13	147.5	4.2	595	5 US-09-673-605A-34	Sequence 34, Appli
14	146	4.2	2843	5 US-09-442-489B-2	Sequence 2, Appli
15	145.5	4.2	1028	6 US-10-179-131-6430	Sequence 6430, Ap
16	143	4.1	1010	5 US-09-914-543-10	Sequence 10, Appli
17	141	4.0	2842	5 US-09-442-489B-7	Sequence 7, Appli
18	141	4.0	2843	4 US-08-681-219A-30	Sequence 30, Appli
19	141	4.0	2843	5 US-09-987-482-1	Sequence 1, Appli
20	141	4.0	2843	7 US-60-389-987-1970	Sequence 1970, Ap
21	139.5	4.0	706	7 US-60-360-039-1865	Sequence 1865, Ap
22	137	3.9	2297	7 US-60-385-568-217	Sequence 217, App
23	136.5	3.9	567	7 US-60-385-568-226	Sequence 226, App
24	136.5	3.9	567	6 US-10-179-131-7530	Sequence 7530, App
25	134	3.8	994	5 US-09-935-625-7547	Sequence 7547, Ap
26	134	3.8	994	5 US-09-935-625-28205	Sequence 28205, A

27	134	3.8	1006	5 US-09-935-625-7546	Sequence 7546, Ap
28	134	3.8	1006	5 US-09-935-625-28204	Sequence 28204, A
29	133.5	3.8	478	5 US-09-445-289B-34	Sequence 34, Appli
30	133.5	3.8	541	7 US-60-360-039-9976	Sequence 9976, Ap
31	133.5	3.8	627	7 US-60-360-039-18499	Sequence 18499, A
32	132	3.8	1014	7 US-60-360-039-1608	Sequence 1608, Ap
33	131.5	3.8	1084	7 US-60-360-039-11155	Sequence 11155, A
34	131.5	3.8	1518	7 US-60-360-039-22243	Sequence 22243, A
35	131.5	3.8	10431	1 PCT-US02-11734-310	Sequence 310, App
36	131	3.8	989	5 US-09-935-625-7548	Sequence 7548, Ap
37	131	3.8	989	5 US-09-935-625-28206	Sequence 28206, Ap
38	131	3.8	1031	5 US-09-935-625-8784	Sequence 8784, Ap
39	131	3.8	1031	5 US-09-935-625-29529	Sequence 29529, A
40	131	3.8	1043	5 US-09-935-625-8783	Sequence 8783, Ap
41	131	3.8	1043	5 US-09-935-625-29528	Sequence 29528, A
42	131	3.8	1332	6 US-10-179-131-9843	Sequence 9843, Ap
43	129	3.7	708	5 US-09-540-209B-6575	Sequence 6575, Ap
44	128	3.7	1029	5 US-09-540-209B-10169	Sequence 10169, A
45	127.5	3.7	623	6 US-10-179-131-5358	Sequence 5358, Ap

ALIGNMENTS

RESULT 1
US-08-961-083-2
Sequence 2, Application US/08961083
GENERAL INFORMATION:

APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961.083

FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 666 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-08-961-083-2

Query Match

Best Local Similarity 100.0%; Score 3484; DB 4; Length 666;

Matches 666; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

1 KIYDKKNLIDLGSESRVYNQANDIPDVKAIYSIDHREFDHGIDTIRILGAPLRN 60

```
|||||
Db 1 KIDNKNQOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHFFPHRGIDITRILGAEFLRN 60
OY LOSNSLOGGSTLTLQOLIKLIFYSTSTSDQITSRKAQAMLAIOLEOKATQOELLTYINK 120
Db 61 LOSNSLOGGSTLTLQOLIKLIFYSTSTSDQITSRKAQAMLAIOLEOKATQOELLTYINK 120
OY 121 VYNSNGNYGQTAQONYGKDLNLSLPOLALLAGMPQAPNOXDYPYSHPEAODRNLVL 180
Db 121 VYNSNGNYGQTAQONYGKDLNLSLPOLALLAGMPQAPNOXDYPYSHPEAODRNLVL 180
OY 181 SEKNNGYISAEQYERKAVNPITDGLQSLKSASNYPAYMDNYLKEVINQVEEFGYMLT 240
Db 181 SEKNNGYISAEQYERKAVNPITDGLQSLKSASNYPAYMDNYLKEVINQVEEFGYMLT 240
OY 241 TGMADVNTNDOEAKHLMIDYNTDEYVAYPPDELOVASTIYDVSNGKVIQOLGARHOSN 300
Db 241 TGMADVNTNDOEAKHLMIDYNTDEYVAYPPDELOVASTIYDVSNGKVIQOLGARHOSN 300
OY 301 VSEGINQAVETNRDMGSTMKPIIDYAPALEGYVDSTATIVHDEPNYPGNTPEVYNMDR 360
Db 301 VSEGINQAVETNRDMGSTMKPIIDYAPALEGYVDSTATIVHDEPNYPGNTPEVYNMDR 360
OY 361 GYRGNTILOALQOOSRNPVAVETLNKGLNRAKTFNLGLGIDPISHYNAISSNTTESD 420
Db 361 GYRGNTILOALQOOSRNPVAVETLNKGLNRAKTFNLGLGIDPISHYNAISSNTTESD 420
OY 421 KKTGASSEKMAAAYAAFANGGTYYKPMYIHKVYVSDGSEKFEFVGRAKETAYAMTD 480
Db 421 KKTGASSEKMAAAYAAFANGGTYYKPMYIHKVYVSDGSEKFEFVGRAKETAYAMTD 480
OY 481 MKMTVLTYGTGRNAYLAMLPOAGKTGTSNYTDEIEHNIKTSQFVAPDELFGATRYKM 540
Db 481 MKMTVLTYGTGRNAYLAMLPOAGKTGTSNYTDEIEHNIKTSQFVAPDELFGATRYKM 540
OY 541 AVMTGVSNRITPLVNGCLTYAAKYRSMATYLLSEGSNPEDEMNIPGGLYRNGEYFKNGAR 600
Db 541 AVMTGVSNRITPLVNGCLTYAAKYRSMATYLLSEGSNPEDEMNIPGGLYRNGEYFKNGAR 600
OY 601 STNNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTNPNNNTQOONTPDOON 660
Db 601 STNNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTNPNNNTQOONTPDOON 660
OY 661 POPAOP 666
Db 661 POPAOP 666

RESULT 2
US-09-769-744A-122
; Sequence 122, Application US/09769744A
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hannlft, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PMC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 122
; LENGTH: 821
; TYPE: prt
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-122
```

```
Query Match 15.2%; Score 529.5; DB 5; Length 821;
Best Local Similarity 27.2%; Pred. No. 4.4e-27;
Matches 203; Conservative 115; Mismatches 270; Indels 159; Gaps 32;

OY 9 LIADLGE-RRVNAQANDIPTDLVKAIVSIEDHFFPHRGIT--DTR-ILGAEFLRLQS 63
Db 115 VIASIESDLRTSSISHQISENLKKAIIATEDEHEKHKGVPRKAVIRATLGKEV-GLGS 173
OY 64 NSLOGGSTLTLQOLIKLIFYSTSTSDQITSRKAQAMLAIOLEOKATQOELLTYINKVYM 123
Db 174 SS--GGSTLTLQOLIKQOVGDA--PTLARKAAEIVDALAREMKNDEILTYLNAAPF 228
OY 124 SNGNYGM-----QTAQONYGKDLNLSLPOLALLAGMPQAPNOXDYPYSHP----- 169
Db 229 GRNKGONIGARAKAAGCIGVDASOLTVQOALAGLPOSPIYSYENTGELKSDEDL 288
OY 170 EAAODRNLVLEKKNNGYISAEQYERKAVNPTI-----TDGLQSL 209
Db 289 EIGLRRAKAVILYSMYRTGALSKEDEYQKDYDLKQDFLPSGTVTGISRDYLYFTTLEAQ 348
OY 210 KSASNYPAYMDNY-LKEVINQVE-----EETGNLTLTGMDVYTNDOEAKH 256
Db 349 ERMVDYLAORDNYSAKELKNEATOKFYRDLAKEIENGGYKITTT-----IDOKIHA 401
OY 257 LMDIYNTDEVVAYPPD--ELOVASTIYDVSNGKVIQOLGARHOSNVSEGINQAVETNR 313
Db 402 MOSA--VADYGYLLDDTGREYGVGNVLMNOTGAILDFVGGRNQOEN--NHAFTKR 456
OY 314 DMGSTMKPIIDYAPALEGYVDSTATIVHDEPNYPGNTPEVYNMDRGYGNTILOALQ 373
Db 457 SPASTKPLLAYGIAIDQGLMS--ETILSNYPFANGNPIMVANSKG-TGMWTLGELN 514
OY 374 QSRNPVAVET--LNKGLNRAKTFNLGLGIDPISHYNAISSNTTESDKKGASSEKM 430
Db 515 YSMNIPAYMTWRMLREKGVV-VKGYMEKMGYEIP-----EGGIESLPM 556
OY 431 AA-----AYAAFANGGTYYKPMYIHKVYVSDGS-EKESNVGTRAMKETAYAM 478
Db 557 GGGLEYVAOHTNGYQYLANGVYHOKHVISKIEAAGGRVYVEQDPVOYISKATITIM 616
OY 479 TDMKTVLTYG-----TGRNAYLAMLPOAGKTGTSNYTDEIEHNIKTSQFVAPDE 529
Db 617 OGILREVLSRVYTTTFKSNLISLNPPLANADWIGKTGTN-QDENMMLMLSTPLFLG- 674
OY 530 LFAGYTRKYSMAVWTGYSN-----RLTPLY-GN-GLTYAAKYRSMATYLLSE 574
Db 675 -WIGHDHNHSLSRAGYSNNSNYMAHLVNAIQASPISMGNERPALDPYVKESEV-LKST 732
OY 575 GSNPEDEMNIP-----EGLYRNGEYFKNGARST-----WNSPAPQOP- 611
Db 733 GQKPEKYSVGEVEYGVTSYTWANKSGAPATSYFAIGSDADYQNMASSTVGSLLPT 792
OY 612 PSTESSSSSDSTSSSTPSTNN 638
Db 793 PSSSSSSSSSDSSNSSTTRPSSRA 819

RESULT 3
US-09-540-209B-8941
; Sequence 8941, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8941
; LENGTH: 789
; TYPE: prt
```

ORGANISM: B. fragilis
US-09-540-209B-8941

Query Match 15.1%; Score 526; DB 5; Length 789;
Best Local Similarity 24.7%; Pred. No. 7.1e-27;
Matches 177; Conservative 118; Mismatches 248; Indels 174; Gaps 25;

```
OY 13 LGSERVNNAQNDIPDILKATVISEDHPRFDHGRDITRILGAFIRN--LGSNSLGGGS 70
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 77 LSKENNVYTSYNELSPNIVALLATEDVRETHSGIDAKALLRAVVRGLLMOKNAGGS 136
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 71 TLTQOIKLTYESTSDOTISRKAQE--AW-LAIOEQATKOEILTYINKVYNSGN 127
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 137 TISQOLAK-QLTDEVARNTLQRLPQKPIEMVIAVKLEKYKKEILSMILNKFDLNN 195
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 128 YQMOTAAONYGKDLNNLSLPOLLALLAGMPQAPNOYDPYSHPEAADRNLVISEKNKG 187
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 196 VQIKTASYTYFGCEPKDLIEQAATLIGMCKNPSLYNPRFNSRSGRRNTVLDQRRKAG 255
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 188 YISAEDEKAVNTPI-----TDGL-----QSLKASNYPRAY-MDNLYK 224
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 256 YTTAEBCDSLQNLPLELVHVRDHEKGLATYERREYLRGVATSKPVRSNTRGMQKFEY 315
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 225 EYI-----NOVEETGYNLTTGMDVYTNVDQAKHLMDI----- 260
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 316 DSDIDEMNPLYGCKEKKKKKSGSNVNYTIDGLKIYTTISHMORVAEAEVEHVEGLQ 375
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 261 -----YNTDEVYAVPDELO----- 275
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 376 LEFEKRRKKAPYSNQLTOEIDRLIDRAVKQTSRYQTMKEAGISEAEIKAFNKPESW 435
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 276 -----VASTI-----VDVSNNGVYINQGLARHOSNSNPFQINQA 308
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 436 SVFTWHGVADDTIMSPDSTRYKHKPLRAGFMSMDPIINGOVAKVGGPNVT--YFOYDMA 492
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 309 VETNRDMSGTMKPTDYAPALEYGV--VDST--ATIYHDEPYNPGTTPVYNDMDRG 362
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 493 MGRGRQVSGTIRPYL-YALAMENGFSPOCETRNVETLILIDENCKPSPKNT-----SKGH 546
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 363 FGN-ITLOYALDOOSRNPVAVETLNKYNLRAKTFELNGLDIDPSIHYSNASISNTTESDK 421
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 547 YGEMVTLKGLANSNMWISAYILMSKLNIPYALARLISHFGVRNKEIQPTVSLCIGPCE--- 603
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 422 KYGASSEKMAAAYAFANGTGYKKPMYIHKVYFSDGSE-KEPSNVSTRAMKETTAIAMD 480
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 604 ---ISVGENVSATYAPANGIRVAPLEVTIKIEDSEGNVLATFSPQMEEVYISASSAYKMLV 660
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 481 MKKTVLTGTGRNA--YLAMLPQAGKTGTSNTDEIEIENHIKTSQVAPDELFAGYTRKY 538
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 661 MLRAVINETGTCARVRRYGTITADMGGKTGTNNRS-----DGMFMGFTPSL 705
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 539 SNAVMTGYSNR---LTPLVGNGLVAAVYRSNM--TYLSE--GSPN-EDMNNIPBG 586
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 706 VSGCWVGGEERDIHEDFTMYTGQASLALPIWTKYMHKVVADQLGYDPKRETFNLPGD 762
    | | | | | : : : | | | | | : | | | | | : : : | | |
```

RESULT 4

US-09-540-209B-9358
; Sequence 9358, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9358
; LENGTH: 790
; TYPE: PRN
; ORGANISM: B. fragilis
US-09-540-209B-9358

Query Match 11.0%; Score 384; DB 5; Length 790;
Best Local Similarity 24.0%; Pred. No. 2.2e-17;
Matches 136; Conservative 102; Mismatches 240; Indels 88; Gaps 21;

```
OY 2 IYNKKQLIADLSEERRVNAQ-----ANDIPTLVAAIVSIEDHPRFDHGRDITRILGA 56
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 66 VYDRNEL---LGARLATQGMRFPPRENTPEVAVCLIEFERORQYHHMGVAPLAIGRA 122
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 57 FLRNLOSNSL-OGSPLJLOOLIKLYFTSTSDOTISRKAQEAMLAIOEQATKOEILT 115
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 123 VYONLKHKRIVSGSSTLTMQITRL---ARNKPRTEKELIEMVATRLFEFRSKKEILS 178
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 116 YTIINKYMSNGNOMTAONTYGKDLNNLSLPOLLALLAGMPAPAPQYDPYSHPEAADR 175
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 179 LYTSHAPFGGNVVGDLAAMRRYFGHSAEELSMASAMLAIVLPNSPAMHLSKROALLK 238
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 176 RNLVISEKNKGYSIAEOYEKAVNTPITDQSLKASNYPAYMDVLYKEVINQVEETG 235
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 239 RNLILHLHKKGLIDTSTYELAISEPLP---QEPILPHIAPHLITDYFYQTRN-----GK 290
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 236 YNLTTGMDVYTNVDQAKHLMDIYNTDEVYAVPDELOVASTIVDSNGKYIAQLGAR 295
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 291 YSVSTIDRGIOQIESLVER--WN---SEFKR--SDIRMLALVIDIRTNQAIAYCGNV 342
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 296 HOSSNVSPGINOAVETNRDMSGTMKPTDYAPALEYGVYDSTIYHDEPYNPGTNPY 355
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 343 HFDKESQGNQVDYIRSPRSTGSLKPLVYAMLQEEELIPN--TLPLDIPVNINGP--TP- 398
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 356 YNMDRGYFNITLOYALDOOSRNPVAVETLNKYNLRAKTFELNGLG---IDPSIHSNAT 412
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 399 QNFNLQPEGAVRSEALIASLNIPSYTLMQRYGVPFHSLKIGLTLNRPSSHGLSL 458
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 413 SSNTTESDKKYGASSEKMAAAYAFANGTGYK--PMY-----IHKVYFSDGSE 459
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 459 I-----LGAEAFTLMDITSAYANMGRSLRPLQFPCTLLDLSISVHRSPFGSAV 509
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 460 KESNVSTRAMKETTYAMTDMMKTYLYTGTGNALAMLPQAGKTGTSNTDEIEIENH 519
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 510 WQTFD---AIKEVNRPEID-WRTIPSMQT---IAW-----KGTGS----- 543
    | | | | | : : : | | | | | : | | | | | : : : | | |
OY 520 KTSQFVAPDELFAGYTRKYSMAVWTG 545
    | | | | | : : : | | | | | : | | | | | : : : | | |
DB 544 ---YGRDAMAVGTPKTAIVGVWVG 565
    | | | | | : : : | | | | | : | | | | | : : : | | |
```

RESULT 5

US-09-914-543-6
; Sequence 6, Application US/09914543
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: LAM, David
; APPLICANT: MATHUR, Eric
; TITLE OF INVENTION: ENDOGLUCANASES
; FILE REFERENCE: DIVER1150-5
; CURRENT APPLICATION NUMBER: US/09/914,543
; CURRENT FILING DATE: 1998-11-22
; PRIOR APPLICATION NUMBER: PCT/US97/08793
; PRIOR FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: US 08/651,572
; PRIOR FILING DATE: 1996-05-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 875
; TYPE: PRN
; ORGANISM: Bankia gouldi
US-09-914-543-6

Query Match 4.6%; Score 162; DB 5; Length 875;
Best Local Similarity 21.1%; Pred. No. 0.018;

; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2283
; TYPE: PRN
; ORGANISM: Staphylococcus epidermidis
PCT-US02-19220-4

Query Match 4.3%; Score 148.5; DB 1; Length 2283;
Best Local Similarity 20.6%; Pred. No. 0.57;
Matches 143; Conservative 79; Mismatches 253; Indels 219; Gaps 31;

QY 63 SNSLOGS-----TLTQOLIKLTFSTSDQTSRKADAMLAILEOKATKQELITYY 117
DB 396 SNVAGGAGAFVTTDSYVATTTYSSTADN-----AAKLNVQPTNPFQDPD 444
QY 118 INKVMNNGN-----YGMQPAON---YGRD-LNLSLPOLALLAGMPQAPN-QYDP 165
DB 445 IN-----YNDGTVMYKYGQTVTRNISPMIAKSGTTNLSLSTASTGATNLOQYQFT 500
QY 166 YSHPEAADRNLVLESEKNOGYISAEOYEKAVNPEIT-DGIQSLKASANNYPAYMDNYLK 224
DB 501 FEYTESAVTQVRVY-DVTGKDIIPKTYSGNVYDVVTIDNOQSALTAK----- 548
QY 225 EYINQVEETGYNLTTGMDVYTNVQEAQKHLMDYNTDEVYAYVDEDLQVASTIVDVS 284
DB 549 -----GYN-----YTSVDS-----SVASTYNDNTKTYKMTN 574
QY 285 NGKVIQLGARHOSNVSEGINOAVEETNRDMGSTMPI-----TDYAPALEYG 332
DB 575 AGOSVYVYTVDKAPVYVG-NQTEYV-----GKTAMPYLVTTDNGTGVNTVGTLPBG 629
QY 333 V-YD-----STATVHDEPNYPGTN-----TPYVNDREG 362
DB 630 LAYDATNSIIGTPKIGGSTVTVSTDOANNKSTTTFTINVDYTAFTVTPIGDQSEV 689
QY 363 FGNIT-LOYALQOSRNVPAVELINKYGLNRAKTF-----LNG-----LGIDYSHYSNA 411
DB 690 YSPISPIKATQDNGNNAVNTVY--TGLPGLFDSTNNTISGTPTNICTSTISIVSTDA 747
QY 412 ISSNTESOKKYGASSEKMAAAYAFANGTYKKPMYIHKYVSDGSEKEFSVGRRAK 471
DB 748 -GONKTTTTFKTEVTRNSMSDS-----VSTSGTQOSQSVSTSKAD 787
QY 472 ETTAYMTDMKTVLTYGTGRNAVYLA MLPQAGKTGTSNYTDEIEHNIKTSQFVADLEF 531
DB 788 SOSASTSGSIVST-----SASTSKSTSVLSDSVASKSLSTSE-- 829
QY 532 AGYTRKYMAVNTGYSNRLTPLVGNGLVYAAKYRSM-----MYLSEG-SNPEDMNIP 584
DB 830 -----SNSVSSSTSTSLVNSQSVSSMSDSASKSTSLDSISNMSSTERS 874
QY 585 EGL-----YRNGEFVFNNGARSTNNSPAPQOPSTESSSSSDS-----STSSQSSST 622
DB 875 ESILSTSTSLKSTSLSLDSLSKSTSGLSKSTSLSTISGSSSTASLSLSDSTSNASTS 934
QY 633 PSTNNSTTNPN-----NNTQOSNTTPOQN 658
DB 935 TSLSESASTSDSTISISNLSANQSASTSKSDSGS 968
RESULT 12
US-09-540-209B-6592
; Sequence 6592, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/540, 209B

; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6592
; LENGTH: 188
; TYPE: PRN
; ORGANISM: B.fragilis
US-09-540-209B-6592

Query Match 4.2%; Score 148; DB 5; Length 188;
Best Local Similarity 27.6%; Pred. No. 0.018;
Matches 48; Conservative 34; Mismatches 76; Indels 16; Gaps 6;

QY 38 EDRFPDHDGIDIRLIGFLRLNQ-SNSLOGSTVLTQOLIKLTFSTSDQTSRKAO 96
DB 4 EDGATFFHGFLEPMARKALYDLKARRAGSTITTMOLVSVFLS---RKNIRAKLE 60
QY 97 EA---WLAIQLEOKATKQELITYYINKVYMSNGNYGQTAQNYGKDLNLSLPOLALL 153
DB 61 EALIVWL-IETRLTSTKERNYEVYLNIVEMGPLVYGVQEAATYFPKRPQSOLTAESIFL 119
QY 154 AGMPQAPNOY-----DPYSHPEAADRNLVLESEKNOGYIS--AEQYEKAVN 199
DB 120 ASITPMPKHFHRSFNNDMLKESLEGYRLITERLVKGIISEVADSIRPEIN 173
RESULT 13
US-09-673-605A-34
; Sequence 34, Application US/09673605A
; GENERAL INFORMATION:
; APPLICANT: The President and Fellows of Harvard College
; TITLE OF INVENTION: REGULATION OF BIOFILM FORMATION
; FILE REFERENCE: 00246/505003
; CURRENT APPLICATION NUMBER: US/09/673, 605A
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/102,870
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 60/083,259
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 595
; TYPE: PRN
; ORGANISM: Escherichia coli
US-09-673-605A-34
Query Match 4.2%; Score 147.5; DB 5; Length 595;
Best Local Similarity 22.7%; Pred. No. 0.099;
Matches 149; Conservative 89; Mismatches 291; Indels 127; Gaps 30;
QY 61 LOSNSLOGSTTLTQOLIKLTFSTSDQTSRKADAMLAI-QLBOKATKQELITYYIN 119
DB 5 INTNSL---SLITQNNINNKQASLSSTIERLS-----SGLRINSARDDAAGQAIANRFTS 56
QY 120 KV-----YMSNGNYGQTAQNYG--KDLNLSLPOLALLAGMPAPNOYDPYSHPEA 172
DB 57 NIKGLTQAARNANDGI-SVAQTTEGALSEINN-NIQRIRBELT--VOASTGTNSDSDLSI 112
QY 173 QDRNLVLESEKNOGYISAEOYEKAVNPTITDQSLKASANNYPAYMDNYLKEVINQVEE 232
DB 113 QDEIKRLEIDR---VSGOTQFNGVYVLA KDQSMKIYQVANDGOTITTTIDKIDSDTLG 169
QY 233 ETGYNLTTG-----MDVYTNVQEAQKHLMDYNTDEVYAYPD--DEL 274
DB 170 LNFENNVNGSGTIANKAATISDLTAAKMDATNT-----ITTTNNALTFASKALDQL 219
QY 275 QVASTIVDVSNGKVIQLGARHOS--NVSEFGINQAVEETNRDMGSTMKPIITDYAPALEYG 333
DB 220 KQGDFTYIRADAQOTATVYVYTNASAGNFSFS--NVSNTSAKAGDVAAASLIPAGQYASGV 278
QY 334 YSTATIVHDEPNYPGTNTPYVNMDRGYFGNTLT--QVALQOSRNVPAVELINKVGLN 390

```

Db      279 YKASGEY-----NFVDVANKKITIGOEAYLTSOGLN!-----TTNDAGCA 319
QY      391 RAKTFPLNL!-GIDYPSIHYSN-AISSNTTESDKKKYKASSEKMAAAYAAPANGTTYK 445
Db      320 TAAT-LDGLFKKADGGSIGFNKTAASYMTGGTTYNFKTGAD-----AGAATAAG- 368
QY      446 PMYIHKVYFSDGSKE--FSNVCGRRAKKEHTAYMYMTDMKTVLTLYGGRNAYLAWLPQAG 503
Db      369 -----SETTASKEVLYLNKVA!-AKQGAVAANGDTSATITTYKSGVOTYAVPAGD 420
QY      504 KTGTSNYDEEIEHNIKTSQFVADFL-FAGYGRKYSMAWMTGY-----SNRLTPL 553
Db      421 GTASAKVADNDVSNATATYTDADGEMTTIGSYTTKTSIDANGSKYTVDSGTSQKAPK 480
QY      554 VGNGLTAAKAYRSMMYTL-----SEGSNPD-WNIPEGLYRNGEEVFKNGARSTWNSP 606
Db      481 VG-----AEVYVANGLLTDTASECTVTKDPLKALDEAIISSIDKFRSSIGAIONRLDS 534
QY      607 APQGPSTESSSSSDSTSSSTSTSTSTNNSTTTNNNNNTQOSNTTPDOONNPQ 662
Db      535 AVTNLNNNTTTLSEQSRIDADYATEVSNNS--KAOITQAGNSVLAANVOPQ 587

```

```

: RESULT 14
: US-09-442-489B-2
: Sequence 2, Application US/09442489B
: GENERAL INFORMATION:
: APPLICANT: Albertsen, Hans
: APPLICANT: Anand, Rakesh
: APPLICANT: Carlson, Mary
: APPLICANT: Groden, Joanna
: APPLICANT: Hedge, Philip John
: APPLICANT: Joslyn, Geoff
: APPLICANT: Kinzler, Kenneth
: APPLICANT: Markham, Alexander Fred
: APPLICANT: Nakamura, Yusuke
: APPLICANT: Thilveris, Andrew
: APPLICANT: Vogelstein, Bert
: APPLICANT: White, Raymond L.
: TITLE OF INVENTION: APC Antibodies
: FILE REFERENCE: 001107.78817
: CURRENT APPLICATION NUMBER: US/09/442.489B
: CURRENT FILING DATE: 1999-11-18
: PRIOR APPLICATION NUMBER: US 08/452,654
: PRIOR FILING DATE: 1995-05-25
: PRIOR APPLICATION NUMBER: US 08/289,548
: PRIOR FILING DATE: 1994-08-12
: PRIOR APPLICATION NUMBER: US 07/741,940
: PRIOR FILING DATE: 1991-08-08
: NUMBER OF SEQ ID NOS: 154
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 2843
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-442-489B-2

```

Query Match	4.28;	Score 146;	DB 5;	Length 2843;
Best Local Similarity	18.88;	Pred. NO. 1.1;		
Matches 130;	Conservative 98;	Mismatches 213;	Indels 250;	Gaps 31;

```

0Y      56  AFLNRLONS-----LOGGSLTQQLIKLTFSTNSDQITSKKAQEAMLAIQLGLE 106
      724  AALNRLMANRPAAKKADANISPGSSLSLHVR-----KO 757
0Y      107  KATKOELLTYINKVYMSNGY-----GMQNAQNYVYKGINLNLSPOLALLAGMPQA 159
      758  KALEAEDLAQHLSTFNNIDLPISKASHRSQMRKQSLYGDYV-----800
0Y      160  PNOYDPLSHPEAAODRRN---LVLSKKKNGGYISAEQYKAVNTPPTDGLSLKSASNV 215

```

```

Db      801  --FDTRHDDNRSNDNFNTGMMNVLSY-----LNTV-----LSSSS 837
Qy      216  PAYMDNTLKEVINOVEETGYNLLTGMIDYTNDOEAKHMDIYNDEVAYPDDE-- 273
Db      838  RGSIDSSSEKDRSLERERG-----IGLGNV-----HPATEMP 870
Qy      274  --LQVASTIVDYSNGKVIAGLQAGARQSSNVSFGINQAVETNRDGN--MKPTLD 324
Db      871  GTSSKRGIGIISTTAQIA--KVMEEVSAHTS-----QEDRSSGTTLHCYTD 917
Qy      325  YAPALEYGVDSFTATVIHDEPVPYPCGNTPEVYMDRGYFGNITLQYA--LOQSRNPAVET 383
Db      918  ERNRLR-----RSSAHTHSTNYN-----TKSESNR-----TCSMPRAKLEYKRS--SND5 963
Qy      384  LNKVGLN-----RAKTFPLNGCID-----YPS-----IHYNSAISSNTTESDK 421
Db      964  LNSVSSNDYGKRGQCKMPESIESYSEDESCKFC5GYQYPADLAIKIHSAHHMDNDGELDT 1022
Qy      422  -----KYGASSEKMAAAYAFANGGITYYKPMTHHVVYSDGSEKFSNVGTRAKKETTA 475
Db      1024  PYNLSKY--SDEQLNCGRQSPSONERMAKPHIIDEIKOSBOROSRN-----QSTY 1074
Qy      476  YNMMDMKKTVITYTGTGNATLAMLPAQAGKTGTSNYDELEIENHIKISQVAPDELFACT 535
Db      1075  YPV-----YTESDDKHLKQPHRGOQECVSPYR 1102
Qy      536  RKYSMAYVTGYSNRLPLVGNGLTVAAKYRSM-----TLYSGSNPEDMNITPE 585
Db      1104  --SR-----GANGSETIRKRVSSNHCINQNVQSILCQEDDYEDDKPTNYSRYSEEGQHEE 1156
Qy      586  GLYRNGEVEFKNGARSTWNSP-----AFQRPSTRESSSSSDSTSQSSSTT--PSTNN 637
Db      1157  ERPNYSILIKYEEKRAHV--DQPIDYSLKAYADIPSSOKOSFSEKSSSGSSQSSKTEHMS5S 1215
Qy      638  STTNPNNNNQSN-----TTPDOONQPOA 664
Db      1216  ENTSTPSSNAKRONQLHPSSAQSRSQGPQAA 1246

```

```

RESULT 15
US-10-179-131-6430
: Sequence 6430, Application US/10179131
: GENERAL INFORMATION:
: APPLICANT: HARE, ROBERTA. S.
: APPLICANT: SHAW, KAREN J.
: APPLICANT: SHIMER JR., GEORGE H.
: APPLICANT: KESSLER, MARCO
: APPLICANT: NOLLING, JORK
: APPLICANT: ZENG, QIANDONG
: APPLICANT: GREENE, JONATHAN R.
: TITLE OF INVENTION: CANDIDA ALBICANS NUCLEIC ACIDS AND POLYPEPTIDES,
: TITLE OF INVENTION: AND USES THEREFOR
: FILE REFERENCE: 2976-4031
: CURRENT APPLICATION NUMBER: US/10/179,131
: CURRENT FILING DATE: 2002-06-21
: NUMBER OF SEQ ID NOS: 10194
: SEQ ID NO 6430
: LENGTH: 1029
: TYPE: PRT
: ORGANISM: Candida albicans
US-10-179-131-6430

```

```

Query Match ..... 4.2%; Score 145.5; DB 6; Length 1029;
Best Local Similarity 18.9%; Pred. No. 0.29;
Matches 141; Conservative 106; Mismatches 233; Indels 267; Gaps 32.

QY 77 IKLFSTSDQIRSRNAQEAAMLAIOEQAKTKQELIYYINKYMSNGNTMGQTAACN 136
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 264 IETATSSVAGDISIAKKCSSKNGPKPIKKTMTMQSMISHSNLL-----KLGATQI 316
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

OY 137 YGKGLNNLSLQALLAGMPQAPMPQYDPYSRPEAAQDRNLVLESMKQGG---YISAEQ 193
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```



```

Db 317 HOKK--REOKLKOQOELKROHOL-NHPDEYFDEALSNKLN---SOYKNTAPTHTNTSVAK 370
QY 194 YEKAVNTPITDGLQSLKASNYPAYMDNYLKEVINQVEETG-----YN- 237
Db 371 LQSLKTPNSSSSASLKLMDKDEAVVPSEQISHDQNDGNGVSGDVESKGERHIFND 430
QY 238 --LTTGMDVYT-----NYDQ-----AQKHLMD-----YN 262
Db 431 EVMQCIADIVYSDPEQRYNDEDEDYDSDDDDDYYQOYERPSNDSLAQSHLYEGDDEIEE 490
QY 263 TDEYVAVPDDE-----LOVASTIVDSNGKVIYAOLGARHOSNVSPGINQA 308
Db 491 ADEEV--EDEDESGEDEDEGGFFLVNKSNN--SNAPIT--LG-OHSASTSTPVAPS 541
QY 309 VETNRBMGSTMKPITD-----YAPALEYGVDS-----TATIVH 342
Db 542 LSRHTD-----ITDDTASITTNKSKYKTIOLEPSTSTINGSDSEDEANPYTSLSH 594
QY 343 -----DEPYNYP-----GTNTPVY-----NMDRGYFGNITLQYALQ 374
Db 595 NVNNDISRGDYDYDNTYTCNPNNSVYASQSPDVVDPENLDMG--SNFDYEF-IEN 651
QY 375 SRVNPAYETLUNKVGLNRAKTFELNGLIDYPSIHYSAISSNTTESDKKYGASSEKMAAY 434
Db 652 NDSIPYVDI-----TFENNSTINNMPISYSP-----SPLSYAISG 688
QY 435 AAFANGTYYKPMYIHKVYFSDGSEKEFSNVGTRAKMETAYMMTDMKTVLTYGGRNA 494
Db 689 GGNKSGVTYVNSPNIYVNVNSNPQOQOQOQAKPKPKTKASPFQLSD----- 734
QY 495 YLAMLPOAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSMAVWTGYSNRLPLV 554
Db 735 -----SEDDSNDSDDDCISG-----LSIG-TRRSSQA----- 761
QY 555 GNGLTAAKAYRSMYTLSEGSNPEDEWNIPEGLYRNGEYFKNGARSTWNSPAPQOPST 614
Db 762 ----LAESYFQSSLSSTQETAPQHHPDAKIEPVAEHVSSINPR--YSSSTISKOPTS 814
QY 615 ESS-----SSSDSSTSOSSSTTPSTNNST----- 639
Db 815 SSSLISQSFEGAGLSSTKELSKSFLGSGTSASTSHDEKTTTIDSSSTGFFQVPPNRD 874
QY 640 -TTNPNNTQOQSNTPPQOQONQOPAQ 665
Db 875 YTPSPDNNTLTRTLSTNFKSKSPLPPO 901

```

Search completed: August 11, 2002, 04:32:42
Job time: 538 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 03:11:49 ; Search time 50.47 Seconds
(without alignments)
1267.991 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KLYDNKNQLIADLGSERRVN.....TQSSNTTPDOQNPNQPAQP 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	3484	100.0	719	2	G95042
2	3475	99.7	719	2	S28036
3	3471	99.6	719	2	S28037
4	3466	99.5	719	2	A42893
5	3466	99.5	719	2	A97913
6	3246	93.2	719	2	S28034
7	3245	93.1	719	2	S28032
8	3243	93.1	719	2	S28031
9	3241	93.0	719	2	S28033
10	3192	91.6	719	2	S28035
11	2820	80.9	608	2	S28036
12	2698.5	77.5	637	2	B42893
13	1657.5	47.6	664	2	G86692
14	1217	34.9	827	2	AD1311
15	1201	34.5	826	2	AD1683
16	1074.5	30.8	886	2	F83862
17	999	28.7	914	2	I40529
18	914	26.2	188	2	S31952
19	817.5	23.5	727	2	B89923
20	813.5	23.3	716	2	S43693
21	802.5	23.0	809	2	F97183
22	739.5	21.2	714	2	AE1353
23	739	21.2	714	2	AE1353
24	724	20.8	714	2	AF1723
25	712	20.4	726	2	F70355
26	698.5	20.0	719	2	A83800
27	670.5	19.2	643	2	E81396
28	656	18.8	966	2	E84053
29	655	18.8	764	2	B97371

30	655	18.8	764	2	AB2589	penicillin-binding
31	649.5	18.6	743	2	D86888	penicillin-binding
32	642.5	18.4	822	2	F83016	penicillin-binding
33	640	18.4	798	2	H81040	penicillin-binding
34	637	18.3	731	2	D95235	penicillin-binding
35	636	18.3	731	2	E98099	penicillin-binding
36	627	18.0	659	2	E64594	penicillin-binding
37	625	17.9	660	2	F71917	penicillin-binding
38	618	17.7	851	2	AF0018	penicillin-binding
39	617	17.7	835	2	B82051	penicillin-binding
40	615	17.7	755	2	A12174	penicillin-binding
41	614	17.6	873	2	B75514	penicillin-binding
42	613.5	17.6	873	2	D84126	penicillin-binding
43	608	17.5	718	2	A13420	penicillin-binding
44	598.5	17.2	624	2	A55220	penicillin-binding
45	598	17.2	885	2	S76357	penicillin-binding

ALIGNMENTS

RESULT 1
G95042 penicillin-binding protein 1A [imported] - Streptococcus pneumoniae (strain TIGR4)
C.Species: Streptococcus pneumoniae
C.Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C.Accession: G95042
R.Retelien, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf
elson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A.Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A.Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A.Reference number: A95000; M01D:21357209; PMID:11463916

A.Accession: G95042

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-719 <KUN>

A.Cross-references: GB:AE005672; PIDN:AAK74536.1; PID:q14971838; GSPDB:GN00164; TIGR:

A.Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0369

C:Superfamily: penicillin-binding protein 1B

Query Match 100.0%; Score 3484; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 4.6e-194;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KLYDNKNQLIADLGSERRVNAQANDIPVDYKRAIYSIEDHREFDHRGIDTIRILGAFLEFN	60
DB	54	KLYDNKNQLIADLGSERRVNAQANDIPVDYKRAIYSIEDHREFDHRGIDTIRILGAFLEFN	113
QY	61	LOSNSLOGSSTLTQOLIKLTFSTSDQTSIRKAOEAMLAQLBOKATKQELIYYINK	120
DB	114	LOSNSLOGSSTLTQOLIKLTFSTSDQTSIRKAOEAMLAQLBOKATKQELIYYINK	173
QY	121	VYMSNGNVMGMQTPAONYKGDNLNLSLPQALLAGMPQAPNDYDPYSHPDEAAQDRRLVL	180
DB	174	VYMSNGNVMGMQTPAONYKGDNLNLSLPQALLAGMPQAPNDYDPYSHPDEAAQDRRLVL	233
QY	181	SEMKNQGYISAOYERKAVNPTTDDGLOSLKSASNPAYADNLTKEVINYVEEETGNLLT	240
DB	234	SEMKNQGYISAOYERKAVNPTTDDGLOSLKSASNPAYADNLTKEVINYVEEETGNLLT	293
QY	241	TGMDEVYTNVDOQAOKHLMIDYNTDEVVAPDDELQVASTIVDVNSGKYTAQAGARHOSN	300
DB	294	TGMDEVYTNVDOQAOKHLMIDYNTDEVVAPDDELQVASTIVDVNSGKYTAQAGARHOSN	353
QY	301	VSEFGINQAVETNRDWSGTTKPTTDYAPALEYGVYDSTATIVHDEPNKPGTNPVYNNMR	360
DB	354	VSEFGINQAVETNRDWSGTTKPTTDYAPALEYGVYDSTATIVHDEPNKPGTNPVYNNMR	413

```
QY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGLIGIDYPSIHYSNAISSNTTESD 420
D 414 GYFGNITLQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGLIGIDYPSIHYSNAISSNTTESD 473
QY 421 KRYGASSEKMAAAVAFAFANGGTYKPMYIHKVVFSDGSEKESFNSVGTAMKETTYAMMTD 480
D 474 KRYGASSEKMAAAVAFAFANGGTYKPMYIHKVVFSDGSEKESFNSVGTAMKETTYAMMTD 533
QY 481 MKKTVLYGTGRNAYLAWLPOAGKTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRKYSM 540
D 534 MKKTVLYGTGRNAYLAWLPOAGKTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRKYSM 593
QY 541 AVMTGYSNRLTPLYGNGLTJVAAKYRSMYTYLSEGSNPEDMNIPGLYRNGEYFVKNGAR 600
D 594 AVMTGYSNRLTPLYGNGLTJVAAKYRSMYTYLSEGSNPEDMNIPGLYRNGEYFVKNGAR 653
QY 601 STWNSPAPQOQPPSTRESSSSSDSSTSSSTPTSTNNSTTNPNNNTQOOSNTTTPDOONON 660
D 654 STWNSPAPQOQPPSTRESSSSSDSSTSSSTPTSTNNSTTNPNNNTQOOSNTTTPDOONON 713
QY 661 POPAQ 666
D 714 POPAQ 719
```

```
RESULT 2
S28038
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 45607) (fragment)
C:Species: Streptococcus pneumoniae
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28038
R:Martin, C.; Sibbold, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of per
A:Reference number: S28031; MUID:93010977
A:Accession: S28038
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67873; NID:947419; PIDN:CAA48073.1; PID:947420
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 99.7%; Score 3475; DB 2; Length 719;
Best Local Similarity 99.7%; Pred. No. 1.5e-193;
Matches 664; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 KIYNKNQNLIDLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
D 54 KIYNKNQNLIDLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
QY 61 LOSNSLOGSSTLTQOLIKLTFFSTSTSDOTSRRKQAEAWLAIOLEOKATKQELITYYINK 120
D 114 LOSNSLOGSSTLTQOLIKLTFFSTSTSDOTSRRKQAEAWLAIOLEOKATKQELITYYINK 173
QY 121 VYMSNGNGMGTAAQNYGKDLNLSLPQALLAGMPQAPNOQDYPYSHPEAQRNLVL 180
D 174 VYMSNGNGMGTAAQNYGKDLNLSLPQALLAGMPQAPNOQDYPYSHPEAQRNLVL 233
QY 181 SEMKNQGTISAEOYEKAVNPITDGLQSLKASNPAYMDNYLKEVINQVEEFTGYNLT 240
D 234 SEMKNQGTISAEOYEKAVNPITDGLQSLKASNPAYMDNYLKEVINQVEEFTGYNLT 293
QY 241 TGMVYTNVDEAQRHLMIDYNTDEYVAYPPDELQVASTIYDVSNKGVIQALGARHSSN 300
D 294 TGMVYTNVDEAQRHLMIDYNTDEYVAYPPDELQVASTIYDVSNKGVIQALGARHSSN 353
QY 301 VSFGINQAVETNRDMGSTMKPTDYAPALEGYVDSTATIVHDEPYNPCTNFTVYWMDR 360
D 354 VSFGINQAVETNRDMGSTMKPTDYAPALEGYVDSTATIVHDEPYNPCTNFTVYWMDR 413
QY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGLIGIDYPSIHYSNAISSNTTESD 420
D 414 GYFGNITLQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGLIGIDYPSIHYSNAISSNTTESD 473
```

```
D 414 GYFGNITLQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGLIGIDYPSIHYSNAISSNTTESD 473
QY 421 KRYGASSEKMAAAVAFAFANGGTYKPMYIHKVVFSDGSEKESFNSVGTAMKETTYAMMTD 480
D 474 KRYGASSEKMAAAVAFAFANGGTYKPMYIHKVVFSDGSEKESFNSVGTAMKETTYAMMTD 533
QY 481 MKKTVLYGTGRNAYLAWLPOAGKTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRKYSM 540
D 534 MKKTVLYGTGRNAYLAWLPOAGKTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRKYSM 593
QY 541 AVMTGYSNRLTPLYGNGLTJVAAKYRSMYTYLSEGSNPEDMNIPGLYRNGEYFVKNGAR 600
D 594 AVMTGYSNRLTPLYGNGLTJVAAKYRSMYTYLSEGSNPEDMNIPGLYRNGEYFVKNGAR 653
QY 601 STWNSPAPQOQPPSTRESSSSSDSSTSSSTPTSTNNSTTNPNNNTQOOSNTTTPDOONON 660
D 654 STWNSPAPQOQPPSTRESSSSSDSSTSSSTPTSTNNSTTNPNNNTQOOSNTTTPDOONON 713
QY 661 POPAQ 666
D 714 POPAQ 719
```

```
RESULT 3
S28037
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 63915) (fragment)
C:Species: Streptococcus pneumoniae
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28037
R:Martin, C.; Sibbold, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28037
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67872; NID:947417; PIDN:CAA48072.1; PID:947418
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 99.6%; Score 3471; DB 2; Length 719;
Best Local Similarity 99.4%; Pred. No. 2.6e-193;
Matches 662; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 KIYNKNQNLIDLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
D 54 KIYNKNQNLIDLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
QY 61 LOSNSLOGSSTLTQOLIKLTFFSTSTSDOTSRRKQAEAWLAIOLEOKATKQELITYYINK 120
D 114 LOSNSLOGSSTLTQOLIKLTFFSTSTSDOTSRRKQAEAWLAIOLEOKATKQELITYYINK 173
QY 121 VYMSNGNGMGTAAQNYGKDLNLSLPQALLAGMPQAPNOQDYPYSHPEAQRNLVL 180
D 174 VYMSNGNGMGTAAQNYGKDLNLSLPQALLAGMPQAPNOQDYPYSHPEAQRNLVL 233
QY 181 SEMKNQGTISAEOYEKAVNPITDGLQSLKASNPAYMDNYLKEVINQVEEFTGYNLT 240
D 234 SEMKNQGTISAEOYEKAVNPITDGLQSLKASNPAYMDNYLKEVINQVEEFTGYNLT 293
QY 241 TGMVYTNVDEAQRHLMIDYNTDEYVAYPPDELQVASTIYDVSNKGVIQALGARHSSN 300
D 294 TGMVYTNVDEAQRHLMIDYNTDEYVAYPPDELQVASTIYDVSNKGVIQALGARHSSN 353
QY 301 VSFGINQAVETNRDMGSTMKPTDYAPALEGYVDSTATIVHDEPYNPCTNFTVYWMDR 360
D 354 VSFGINQAVETNRDMGSTMKPTDYAPALEGYVDSTATIVHDEPYNPCTNFTVYWMDR 413
QY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGLIGIDYPSIHYSNAISSNTTESD 420
D 414 GYFGNITLQYALQOOSRNPVAVETLNKVGILNRAKTFPLNGLIGIDYPSIHYSNAISSNTTESD 473
```



```
|||||
Db 414 GYFNITLQYALQOOSRNPVAVETLNKGLNRAKTFNLGLGIDYPSIHNSAISSNTTSD 473
OY 421 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKESNVGTAMKETTAYMTD 480
Db 474 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKESNVGTAMKETTAYMTD 533
OY 481 MKKTYLTYGTGRNAVLA LPAQAGKTGTSNYTDEEIEHNIKTQSFAVAPDELFAGYTRKYSM 540
Db 534 MKKTYLTYGTGRNAVLA LPAQAGKTGTSNYTDEEIEHNIKTQSFAVAPDELFAGYTRKYSM 593
OY 541 AVMTGYSNRLTPVNGCLTVAAKYRSMYTLSEGSNPEDNNIPGGLRNGEFPFKNGAR 600
Db 594 AVMTGYSNRLTPVNGCLTVAAKYRSMYTLSEGSNPEDNNIPGGLRNGEFPFKNGAR 653
OY 601 STWNSPAPQOPSTRESSSSSSDSSSTPSTPSTNNSTTNPNNNTQOQSTTTPDOQON 660
Db 654 STWNSPAPQOPSTRESSSSSSDSSSTPSTPSTNNSTTNPNNNTQOQSTTTPDOQON 713
OY 661 POPAOP 666
Db 714 POPAOP 719
```

```
RESULT 6
S28034
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 56742) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28034
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of per
A:Reference number: S28031; MUID:93010977
A:Accession: S28034
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67869; NID:947411; PIDN:CAA48069.1; PID:947412
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 93.28; Score 3246; DB 2; Length 719;
Best Local Similarity 92.08; Pred. No. 2.8e-180;
Matches 613; Conservative 31; Mismatches 22; Indels 0; Gaps 0;
```

```
OY 1 KIYNKNOLINDLSERRVNAQANDIPTDLKATVSIEDHFPDRGIDTIRILGAFLRN 60
Db 54 KIYNKNOLINDLSERRVNAQANDIPTDLKATVSIEDHFPDRGIDTIRILGAFLRN 113
OY 61 LOSNSLOGGSTLTQOLIKLTFEYSTSDPTISRKAQEWMLAIOLEOKATKOEILTYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLTFEYSTSDPTISRKAQEWMLAIOLEOKATKOEILTYINK 173
OY 121 VYMSNGNGMGTAAONYYGKDLNLSLPOLALLAGMPAPNOYPSHPEAODRRNLVL 180
Db 174 VYMSNGNGMGTAAONYYGKDLNLSLPOLALLAGMPAPNOYPSHPEAODRRNLVL 233
OY 181 SEMNNOGITSAEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVDEETGYNLT 240
Db 234 SEMNNOGITSAEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVDEETGYNLT 293
OY 241 TGMVYTNVDEQAKHLMDIYNTEDEYVAYPDDELQVASTIYDVNSGKVIQALGARHOSSN 300
Db 294 TGMVYTNVDEQAKHLMDIYNTEDEYVAYPDDELQVASTIYDVNSGKVIQALGARHOSSN 353
OY 301 VSEGINAVETNRDWSGTMKPTDYAPALEYGYVDSTATIVHDEPYNPGTNTPEYMMDR 360
Db 354 VSEGINAVETNRDWSGTMKPTDYAPALEYGYVDSTATIVHDEPYNPGTNTPEYMMDR 413
OY 361 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKTFNLGLGIDYPSIHNSAISSNTTSD 420
Db 414 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKTFNLGLGIDYPSIHNSAISSNTTSD 473
```

```
OY 421 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKESNVGTAMKETTAYMTD 480
Db 474 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKESNVGTAMKETTAYMTD 533
OY 481 MKKTYLTYGTGRNAVLA LPAQAGKTGTSNYTDEEIEHNIKTQSFAVAPDELFAGYTRKYSM 540
Db 534 MKKTYLTYGTGRNAVLA LPAQAGKTGTSNYTDEEIEHNIKTQSFAVAPDELFAGYTRKYSM 593
OY 541 AVMTGYSNRLTPVNGCLTVAAKYRSMYTLSEGSNPEDNNIPGGLRNGEFPFKNGAR 600
Db 594 AVMTGYSNRLTPVNGCLTVAAKYRSMYTLSEGSNPEDNNIPGGLRNGEFPFKNGAR 653
OY 601 STWNSPAPQOPSTRESSSSSSDSSSTPSTPSTNNSTTNPNNNTQOQSTTTPDOQON 660
Db 654 PIMTEPSTQOQSTRESSSSSSDSSSTPSTPSTNNSTTNPNNNTQOQSTTTPDOQON 713
OY 661 POPAOP 666
Db 714 POPAOP 719
```

```
RESULT 7
S28032
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 681) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28032
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28032
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67866; NID:947407; PIDN:CAA48066.1; PID:947408
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 93.18; Score 3245; DB 2; Length 719;
Best Local Similarity 92.08; Pred. No. 3.2e-180;
Matches 613; Conservative 35; Mismatches 18; Indels 0; Gaps 0;
```

```
OY 1 KIYNKNOLINDLSERRVNAQANDIPTDLKATVSIEDHFPDRGIDTIRILGAFLRN 60
Db 54 KIYNKNOLINDLSERRVNAQANDIPTDLKATVSIEDHFPDRGIDTIRILGAFLRN 113
OY 61 LOSNSLOGGSTLTQOLIKLTFEYSTSDPTISRKAQEWMLAIOLEOKATKOEILTYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLTFEYSTSDPTISRKAQEWMLAIOLEOKATKOEILTYINK 173
OY 121 VYMSNGNGMGTAAONYYGKDLNLSLPOLALLAGMPAPNOYPSHPEAODRRNLVL 180
Db 174 VYMSNGNGMGTAAONYYGKDLNLSLPOLALLAGMPAPNOYPSHPEAODRRNLVL 233
OY 181 SEMNNOGITSAEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVDEETGYNLT 240
Db 234 SEMNNOGITSAEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVDEETGYNLT 293
OY 241 TGMVYTNVDEQAKHLMDIYNTEDEYVAYPDDELQVASTIYDVNSGKVIQALGARHOSSN 300
Db 294 TGMVYTNVDEQAKHLMDIYNTEDEYVAYPDDELQVASTIYDVNSGKVIQALGARHOSSN 353
OY 301 VSEGINAVETNRDWSGTMKPTDYAPALEYGYVDSTATIVHDEPYNPGTNTPEYMMDR 360
Db 354 VSEGINAVETNRDWSGTMKPTDYAPALEYGYVDSTATIVHDEPYNPGTNTPEYMMDR 413
OY 361 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKTFNLGLGIDYPSIHNSAISSNTTSD 420
Db 414 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKTFNLGLGIDYPSIHNSAISSNTTSD 473
OY 421 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKESNVGTAMKETTAYMTD 480
```


Db	534	MMKTYLAVGTGGGATLPMLAQAGKGTGSNTYTDDELEKHKKNTGYAPDEMFQYTRKYSM	593
Qy	541	AVMTGYSNRRLPLPVNGLTLYAAKYRSMATYISSEGSNPEPDNIPEGLYRNGEYFKNGAR	600
Db	594	AVMTGYSNRRLPLPYVDGGLYAAKYRSMITLYISEDTHPEPDMMPGGLFRNGEYFEVKNGAR	653
Qy	601	STWNSPAQQPSTRESSSSSDSTSQSSSTTPSTNNSTTTPNNNTQQSNTTPOQONN	660
Db	654	STWSSPAQQPSTRESSSSSDSTSQSSSTTPSTNNSTTTPNNNTQQSNTTTPQOQNN	713
Qy	661	POPAPQ 666	
Db	714	POPAPQ 719	

RESULT 10
S28035
penicillin-binding protein 1A - Streptococcus pneumoniae (strain 2039) (fragment)
C:Species: Streptococcus pneumoniae
A:Variety: strain 2039
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28035
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant strains of *Streptococcus pneumoniae*
A:Reference number: S28031; MUID:93010977
A:Accession: S28035
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67870; NTD:g47413; PIDN:CAA48070.1; PID:g47414
A:Experimental source: strain 2039
A:Superfamily: penicillin-binding protein 1B
A:Keywords: antibiotic resistance; penicillin resistance

Query Match	91.68;	Score 3192;	DB 2;	Length 719;
Best Local Similarly	90.48;	Pred. No. 3.8e-177;		
Matches 602; Conservative	41;	Mismatches 23;	Indels 0;	Gaps 0;

Qy	1	KIYNNKOLLJADLCSERRVNAQANDIPDULVKAIVSIEDHRFPDHRGSDITRILIGAFNR	60
Db	54	KIFPSKNNLJADLCSERRVNAQANEITDULVKAIVSIEDHRFPDHRGVDIRILIGAFNR	113
Qy	61	LÖSNLSOGGSLTLQOOLIKLTYFSTSTSDQTSRKAQAMLAIDLEKATQOEILTYTYNK	120
Db	114	LÖNNLSOGGSLTLQOOLIKLTYFSTSTSDQTSRKAQAMLAIDLEKATQOEILTYTYNK	173
Qy	121	VYMSNGNVMQMTAAQANYGYKDLNMLSPOLATLAGMPOARNOQDPSHREPAQDRRLVL	180
Db	174	VYMSNGNVMQMTAAQANYGYKDLNMLSPOLATLAGMPOARNOQDPSHREPAQDRRLVL	233
Qy	181	SEKNNÖGYSIAEÖYEKAVNPIITDGLDLSKSASNYPRAYMDNYLKEVYNÖYBEETYNLTT	240
Db	234	SEKMDÖGYSIAEÖYEKAVNPIITDGLDLSKSASNYPRAYMDNYLKEVYNÖYBEETYNLTT	293
Qy	241	TGMÖVUTYVWDDEAOKHLMIDINTDEYAAVPDDELOVASTIDVNSNGKVIJQLAGRHOSN	300
Db	294	TGMÖVUTYVWDDEAOKHLMIDINTDQYVYSPDDLOVASTVVDVNSNGKVIJQLAGRHOSN	353
Qy	301	VSEGINQAVENIRBOWGSTMKPIITDYAPALEGYVDSTATYVHDEPPYVPGCTNPIPVNMDR	360
Db	354	VSEGINQAVENIRBOWGSTMKPIITDYAPALEGYVDSTATYVHDEPPYVPGCTNPIPVNMDR	413
Qy	361	GYPENITLOYLQOOSRNVPAEITLNKQGLNNAKFTLNGIGIDIPYSIHSAISSNTTESD	420
Db	414	VYFENITQYVALQOOSRNVPAEITLNKQGLDRAKFTLNGIGIDIPYSIMYAHSAISSNTTESN	473
Qy	421	KKYGAASSEKMAAATAPANGSTYKPMYIHKHYVESDGESEKFSNVGTRAKETTAYMTD	480
Db	474	KÖYGAASSEKMAAATAAPANGSTYKPMYINKIYVESDSEKFEPSAGTRAKETTAYMTE	533

[illegible]

RESULT 11
S28036
penicillin-binding protein Ia - Streptococcus pneumoniae (strain 8250) (fragment)
C.Species: Streptococcus pneumoniae
C.Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C.Accession: S28036
R.Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A.Title: Relatedness of penicillin-binding protein Ia genes from different clones of
A.Reference number: S28031; MUID:93010977
A.Accession: S28036
A.Status: translation not shown
A.Molecule type: DNA
A.Residues: 1-608 <MAR>
A.Cross-references: EMBL:X67871; NID:g47415; PIDN:CAA48071.1; PID:g47416
C.Superfamily: penicillin-binding protein IB

Query Match	80.98	Score 2820	DB 2	Length 608
Best Local Similarity	96.88	Pred. No. 1e-155		
Matches 537	Conservative 10	Mismatches 8	Indels 0	Gaps 0

Oy	1	KIYDKNOLJIDLSEFRRVNAQANDIPDTOLYKAVSIJEDHFFPHRGIDTRIIGAFLRN	60
Dd	54	KIYDKNOLJIDLSEFRRVNAQANDIPDLKAVSIJEDHFFPHRGIDTRIIGAFLRN	113
Oy	61	LOSNSLOGSTLTOOLIKLYFSTSTSDQTSIRKAOEAMLAIOLEOKATYOEILTYIYINK	120
Dd	114	LOSNSLOGSTLTOOLIKLYFSTSTSDQTSIRKAOEAMLAIOLEOKATYOEILTYIYINK	173
Oy	121	YVMSNGNGMOTAAONNYGKDLNLSI.POLALLAGMOPARQOVPYSHPEAODRRMLVL	180
Dd	174	YVMSNGNGMOTAAONNYGKDLNLSI.POLALLAGMOPARQOVPYSHPEAODRRMLVL	233
Oy	181	SEMKNOGYISAEQYEKAVNPITDTGLOSLSASNPYPAYMDNYLKEVINQOEETGYNLLT	240
Dd	234	SEMKNOGYISAEQYEKAVNPITDTGLOSLSASNPYPAYMDNYLKEVINQOEETGYNLLT	293
Oy	241	TGMQVYTNVDEQAKHLMDIYNDEBYAAYPPDELOQVASTIVDSNGKVINQOLARHOSSN	300
Dd	294	TGMQVYTNVDEQAKHLMDIYNDEBYAAYPPDELOQVASTIVDSNGKVINQOLARHOSSN	353
Oy	301	VSFQINQAVETNRBMGSTMKPIITYAPALLEGYVDSATIVVHDEPNYPCTNTPYVWMDR	360
Dd	354	VSFQINQAVETNRBMGSTMKPIITYAPALLEGYVDSATIVVHDEPNYPCTNTPYVWMDR	413
Oy	361	GYFGNITLOVALOOSRNPVAVETLNKVGLENKAKFTLNGLGIDPYSIHSNAISSNTTESD	420
Dd	414	GYFGNITLOVALOOSRNPVAVETLNKVGLENKAKFTLNGLGIDPYSIHSNAISSNTTESD	473
Oy	421	KKYGSSSEKMAAAYAAPANGSTGYKKPMYIHKVYVSDGSEKESVNGVRRAKKEITAYVMTD	480
Dd	474	KKYGSSSEKMAAAYAAPANGSTGYKKPMYIHKVYVSDGSEKESVNGVRRAKKEATAYMTE	533
Oy	481	MKKTIVLYYGTGRNMYLAMLPOAGKTGTSNYTDEEINHIISQFVAVDELFCAGYTRKYS	540

Db 534 MKKTVLTVGIGRGAYLPMLPQAGKTGTSNYTDEIEIKYIKNTGYVAPDEMEVGYTRKYAM 593
OY 541 AVMTGYSNRLPLVG 555
Db 594 AVMTGYSNRLPLIG 608

RESULT 12

penicillin-binding protein 1 - Streptococcus oralis
C:Species: Streptococcus oralis
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 04-Mar-2000
C:Accession: B42893
R:Martin, C.; Briese, T.; Hakenbeck, R.
J. Bacteriol. 174, 4517-4523, 1992
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus oralis
A:Reference number: A42893; MUID:92325042
A:Accession: B42893
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-637 <MAR>
A:Cross-references: GB:M90528; NID:9153769; PIDN:AAA26958.1; PID:9153771
C:Superfamily: penicillin-binding protein 1B

Query Match 77.5%; Score 2698.5; DB 2; Length 637;
Best Local Similarity 87.5%; Pred. No. 1.2e-148;
Matches 511; Conservative 39; Mismatches 33; Indels 1; Gaps 1;

OY 1 KIYDNKNOIADLGSERRVNAOANDIPTDLYKAIVSIEDHREFDHRGIDTIRILGAFLEFN 60
Db 54 KIYDNDELADLGSEKRYNAOANEIPDLYNAIVSIEDHREFNRRGIDTIRILGATLEFN 113
OY 61 LQ-SNSLOGGSTLTQOLIKLTFYFSTSDPTISRKAQEAMLAIOLEOKATKOEILTYYN 119
Db 114 LRGGGGLGASTLTQOLIKLTFYFSTSDPTISRKAQEAMLAIOLEOKATKOEILTYYN 173
OY 120 KYMNSNGYGMOTAAQNTYTKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAQQDRNLV 179
Db 174 KYMNSNGYGMOTAAQNTYTKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAQQDRNLV 233
OY 180 LSEMNQGYIASOEYKAVNPTITDGLQSLKASNPAYMDNYLKEVYNQVEEETGYNLT 239
Db 234 LSEMGQGYITAEYKAIPTITDGLQSLKASNPAYMDNYLKEVYNQVEEETGYNLT 293
OY 240 TTGMVYTNVDQEAOKHLMDIYNTDEYVAYPPDELOVASTYVDSNGKYIAQLGARHSS 299
Db 234 TTGMVYTNVDQEAOKHLMDIYNTDEYVAYPPDELOVASTYVDSNGKYIAQLGARHSS 353
OY 300 NVSEGINOAVETNRDMGSTMKPTITDYAPALEYGYVDSTATIVHDEPYNPGTNPVYVMD 359
Db 354 NVSEGINOAVETNRDMGSTMKPTITDYAPALEYGYVDSTATIVHDEPYNPGTNPVYVMD 413
OY 360 RGYEENITLOVALQOQRNPAVETLNKYGILNRAKTFLENGIDYPSIHNSAISNTTES 419
Db 414 KSYFENITLOVALQOQRNPAVETLNKYGILNRAKTFLENGIDYPSIHNSAISNTTES 473
OY 420 DKKYASSEKMAAAAYAFANGSTYKPMYIHKVFSDSGESEKFSNVGTGRAMETAYMMT 479
Db 474 DKKYASSEKMAAAAYAFANGSTYKPMYIHKVFSDSGESEKFSNVGTGRAMETAYMMT 533
OY 480 DMKTVLTVGIGRGAYLPMLPQAGKTGTSNYTDEIEIKYIKNTGYVAPDEMEVGYTRKY 539
Db 534 DMKTVLTVGIGRGAYLPMLPQAGKTGTSNYTDEIEIKYIKNTGYVAPDEMEVGYTRKY 593
OY 540 MAVMTGYSNRLPLVGNGITVAAKYRSMYTLSEGSNPEDMNI 583
Db 594 MAVMTGYSNRLPLVGNGITVAAKYRSMYTLSEGSNPEDMNI 637

RESULT 13
G86692

penicillin-binding protein 1A [imported] - Lactococcus lactis subsp. lactis (strain I
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86692
R:Boletín, A.; Wincker, P.; Mager, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Eh
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86692
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-664 <STO>
A:Cross-references: GB:AF005176; PID:912723428; PIDN:AAK04641.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ponA
C:Superfamily: penicillin-binding protein 1B

Query Match 47.6%; Score 1657.5; DB 2; Length 664;
Best Local Similarity 52.1%; Pred. No. 2.6e-88;
Matches 333; Conservative 105; Mismatches 166; Indels 35; Gaps 6;

OY 1 KIYDNKNOIADLGSERRVNAOANDIPTDLYKAIVSIEDHREFDHRGIDTIRILGAFLEFN 60
Db 49 KYIDSEKRYVATLGAQQRNLYKTDNIPVLYNAIVSIEDHREFNRRGIDTIRILGAFVFN 108
OY 61 LQ-SNSLOGGSTLTQOLIKLTFYFSTSDPTISRKAQEAMLAIOLEOKATKOEILTYYN 120
Db 109 LRGGGGLGASTLTQOLIKLTFYFSTSDPTISRKAQEAMLAIOLEOKATKOEILTYYN 168
OY 121 KYMNSNGYGMOTAAQNTYTKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAQQDRNLV 180
Db 169 VMANGYGYGMOTAAQNTYTKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAQQDRNLV 228
OY 181 SEMKNQGYIASOEYKAVNPTITDGLQSLKASNPAYMDNYLKEVYNQVEEETGYNLT 240
Db 229 RAMRRYDKTITAEKKAATLPIDDLQPLKQSVTIPSYDNFLKQALQAKTLAGDITLT 288
OY 241 TTGMVYTNVDQEAOKHLMDIYNTDEYVAYPPDELOVASTYVDSNGKYIAQLGARHSS 300
Db 289 EGAKIYTTLTDTAQNLYNIVNTGNITPTDPMQVASTVDTQKGAVALQIGRRQPSN 348
OY 301 VSEGINOAVETNRDMGSTMKPTITDYAPALEYGYVDSTATIVHDEPYNPGTNPVYVMD 360
Db 349 VTFGNQAVQTDQDMGSTMKPTITDYAPALEYGYVDSTATIVHDEPYNPGTNPVYVMD 407
OY 361 RGYEENITLOVALQOQRNPAVETLNKYGILNRAKTFLENGIDYPSIHNSAISNTTES 420
Db 408 TFGSMYVASALASRNIPAVKTLNVLGDNSSKTFVNGIGITLDPLETSMALSSNS- 463
OY 421 DKKYASSEKMAAAAYAFANGSTYKPMYIHKVFSDSGESEKFSNVGTGRAMETAYMMT 480
Db 464 DKKYASSEKMAAAAYAFANGSTYKPMYIHKVFSDSGESEKFSNVGTGRAMETAYMMT 523
OY 481 MKKTVLTVGIGRGAYLPMLPQAGKTGTSNYTDEIEIKYIKNTGYVAPDEMEVGYTRKY 532
Db 524 ILQSVTLTVLSESVSYAIVPGLAAAGKTGTSNYTDEIEIKYIKNTGYVAPDEMEVGYTRKY 583
OY 533 GYTRKYSMAVMYGTYSNRLPLVGNGITVAAKYRSMYTLSEGSNPEDMNIPEGYTRNG 591
Db 584 GTTPQYSMAVMYGTYSNRLPLVGNGITVAAKYRSMYTLSEGSNPEDMNIPEGYTRNG 643
OY 592 EFVFRNGARSTWNSPAPQPPSTESSSSSSSSSTSSSS 630
Db 644 TALVK-----TSSSGQTISQSSA 661

RESULT 14
AD1311
penicillin-binding protein 2A homolog pbpa [imported] - Listeria monocytogenes (strain
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: ADJ311
R:Glasser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baqerou, F.; Berche, P.; Blocker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefelt, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tlieretz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A.Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: ADJ311
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-827 <GLA>
A:Cross-references: GB:NC_003210; PIDW:CAC9970.1; PID:g16411345; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: pbpA

Query Match	34.9%	Score 1217	DB 2	Length 827
Best Local Similarity	40.3%	Pred. No. 1.2e-62		
Matches 268	Conservative 133	Mismatches 210	Indels 54	Gaps 14

Query Match	Similarity	34.9%	Score 1217	DB 2	Length 827
Best Local Similarity	40.3%	Pred. No. 1.2e-62			
Matches 268	Conservative 133	Mismatches 210	Indels 54	Gaps 14	
QY	1	KIYDKNQLLIADLGSERRVNAQANIPIDILVAIYSIEDHREDFHGRGIDTIRILGAPLRN	60		
Db	79	KLLDDGKVFPAEVEGERREYETEKRIPEPLKNAILLTTEEDARYEHGDIDPILGCAVIAN	138		
QY	61	LOS-N-SLOGSTLTLQOLIKLYFTSTSDQTSIRKAOEMALAIOLKQATKOELLTYIAN	119		
Db	139	LTDGAGAGASTLSQOITIKMSYL--DYNTKTLARKAOEMALAIOLKEKSKMDLEITYVN	196		
QY	120	KVYMGNGVGMQTOAONYYGKDNLNLSPQLALLGMPQAPMOYPPYSPHEAQRRLIV	179		
Db	197	KVYMSDRHMGQTAEEHFGKNVKLLTAQALLAGMPOSPNNYPPYEPHEAKRRPOV	256		
QY	180	LSEMNQGYISAEQYERKAVNPETDGLQSLKSASVYPAVMNYLKEVINQVEERGNYL	239		
Db	257	LTNMTTHKIRTEEMTEAKPTPIITGLSKSKDKREKIRKYDSYVQVLSLPKE--YDY	314		
QY	240	TTGMDVYTNVDQEAQHLMIDYINDEYVAYPDDLEQVASTIVDSNGVIAQLGARHSS	299		
Db	315	RDGLTIHIALDRDAQEYETKMLNTEIYNFTDDEMQAGIVLQDTKGTGRVQALGGGRQ--	372		
QY	300	NVSPFINQAVETNRNMGSTMKPITDYAPALEGYVDSTATIYHDEPYNPGTINPFYVMD	359		
Db	373	KYTRGYNATQYKRSGVGTMKPDIADYGPAFEYLD--STAHLLEDPEPTYSG--GPIINMD	430		
QY	360	RGYFENITLOYALQOOSRVNPAVETLNKGLIRAKFTFLNGLDIYDSIH--YSNATISSMT	417		
Db	431	FGYKPIYRQALVOSRNIPALKTIOAVGLDKSEFVKNLGIITDYGCGNVESNATGANS--	489		
QY	418	ESDKKYGASSEKMAAAYAFANGFTYKPMYIHKVFEFS--SEKFSNVGTRAMKETAY	476		
Db	490	-----SNPQMGAYAFACNKGITGNKPHVYTKIVLSDGQTEIDTEPQSTVAMKESYAY	542		
QY	477	MATDMKTVLTYGTGRNAYLAMLPOAGKTGTSNYTDEEIEHHIKTSQFVAP---DEL	531		
Db	543	MYSVLKDLVLTIGTGTSAAVPGVPAAGTGTGINPPE-----FTSKYYYPGGAARDSMF	596		
QY	532	AGYTRKYMAAVWTGGSNNLTPLVGNGLTVAAKYVRSMNTYLSSEGSNPEPDMMINPEGLYNG	591		
Db	597	AGYTTNYSIAWVTGDDDKKKYVASSEOKIAORMFESKMAHNASAGCTTADFPKMPSVV---	653		
QY	592	EVEFKNGARITWNSPAP---QQPPSTSSSSSSSSSTSSQ--SSSTPPSTNNSTTTNN	645		
Db	654	-----SVPLKGSNPIARAAQGTSSDKVSYELFLSCTAPPTASTPDEDEKK	699		
QY	646	NTQOS 650			
Db	700	KAEBA 704			

AD1683
penicillin-binding protein 2A homolog pbppa [imported] - *Listeria innocua* (strain Cl1p
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence-revision 27-Nov-2001 #text-change 27-Nov-2001
C:Accession: AD1683
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Blöc-
k, C.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fsihl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.;
Ok, C.; Schlueter, T.; Simoes, N.; Tlerraz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077, MUID:21537279, PMID:11679669
A:Accession: AD1683
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-826 <GFA>
A:Cross-references: GB:AL592023; PIDN:CAC97236.1; PID:g16414507, GSPDB:GN00178
A:Experimental source: strain Cl1p1262
A:Genetics:
A:Gene: pbppa

[illegible]

Db 709 AEDKKTEEEKKEEAK 727

Search completed: August 11, 2002, 04:24:15
Job time: 4346 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 04:25:34 ; Search time 36.95 Seconds

(without alignments)
697.895 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 K1YDNKNQLADIGSERV.....TQOSNTTPDQGNQNPQAP 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	719	1	PPPA_STRPN
2	2698.5	77.5	637	1	PPPA_STROR
3	999	28.7	914	1	PPPA_BACSU
4	739	21.2	714	1	PPPE_BACSU
5	712	20.4	726	1	PPPA_AQUAE
6	680.5	19.5	798	1	PPPA_NEITL
7	642.5	18.4	822	1	PPPA_PSEAE
8	640	18.4	798	1	PPPA_NEIMA
9	640	18.4	798	1	PPPA_NEIMA
10	635.5	18.2	797	1	PPPA_NEIGO
11	633	18.2	798	1	PPPA_NEICI
12	617	17.7	825	1	PPPA_VIBCH
13	598.5	17.0	624	1	PPPD_BACSU
14	591.5	17.0	853	1	PPPA_HAEIN
15	584	16.8	777	1	PPPV_VIBCH
16	567	16.3	850	1	PPPA_ECOLI
17	526	15.1	844	1	PPPB_ECOLI
18	525.5	15.1	781	1	PPPB_HAEIN
19	502.5	14.4	787	1	PPPA_RICPR
20	490.5	14.1	809	1	PPPA_XYUFA
21	472	13.5	760	1	PPPB_BUCAI
22	450.5	12.9	770	1	PPPC_ECOLI
23	312.5	9.0	207	1	TRG_ALCEU
24	207	5.9	716	1	PPPB_BACSU
25	206	5.9	233	1	MTGA_NEIGO
26	206	5.9	645	1	SPSD_BACSU
27	194.5	5.6	224	1	MTGA_ACICA
28	193.5	5.6	598	1	FTSI_MESVI
29	192	5.5	233	1	MTGA_NEIMA
30	185.5	5.3	242	1	MTGA_KLEPM
31	185.5	5.3	242	1	MTGA_ECOLI
32	181	5.2	246	1	MTGA_HAEIN
33	164	4.7	588	1	FTSI_ECOLI

34	147	4.2	1861	1	APU_THETU	P36536 t amylopull
35	142.5	4.1	750	1	PPBX_STRPN	P14677 streptococ
36	142	4.1	610	1	FTSI_HAEIN	P45059 haemophilus
37	141	4.0	716	1	YOGF_BACSU	P54488 bacillus su
38	140.5	4.0	2843	1	APC_HUMAN	P25054 homo sapien
39	140.5	4.0	516	1	P54_ENTFC	P13692 enterococcu
40	139.5	4.0	706	1	PLB2_YEAST	Q03674 saccharomyc
41	138	4.0	1337	1	DEXT_STRDO	P39653 streptococc
42	136.5	3.9	567	1	CHT3_CANAL	P40954 candida alb
43	136	3.9	579	1	FTSI_BUCAI	P57317 buchnera ap
44	135.5	3.9	1007	1	V741_CHLMG	O9p16 chlamydia m
45	135	3.9	1365	1	GTF5_STRDO	P29336 streptococc

ALIGNMENTS

RESULT ID	1	PPPA_STRPN	STANDARD:	PRT:	719 AA.
AC	004707	PPPA_STRPN			
DT	01-OCT-1994 (Rel. 30, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	penicillin-binding protein 1A (PBP-1A) (Exported protein 2).				
GN	POMA OR EXP2 OR SP0369.				
OS	Streptococcus pneumoniae.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				
OX	Streptococcus.				
NCBI_TaxID=1313;					
RP	SEQUENCE FROM N.A.				
RC	STRAIN=45607, AND 63915;				
RC	MEDLINE=93010977; PubMed=1396576;				
RA	Martin C., Sibold C., Hakenbeck R.;				
RT	"Relatedness of penicillin-binding protein 1a genes from different				
RT	clones of penicillin-resistant Streptococcus pneumoniae isolated in				
RT	South Africa and Spain".;				
RL	EMBO J. 11:3831-3836(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=R6;				
RC	MEDLINE=92325042; PubMed=1624444;				
RA	Martin C., Briese T., Hakenbeck R.;				
RT	"Nucleotide sequences of genes encoding penicillin-binding proteins				
RT	from Streptococcus pneumoniae and Streptococcus oralis with high				
RT	homology to Escherichia coli penicillin-binding proteins 1a and 1b.";				
RL	J. Bacteriol. 174:4517-4523(1992).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=TRIG4;				
RC	MEDLINE=21357209; PubMed=11463916;				
RA	Tetzelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,				
RA	Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,				
RA	Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,				
RA	Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,				
RA	Holtzapfel E., Knouri H., Wolf A.M., Ueberback T.R., Hansen C.L.,				
RA	McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,				
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,				
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;				
RT	"Complete genome sequence of a virulent isolate of Streptococcus				
RT	pneumoniae.";				
RL	Science 293:498-506(2001).				
RP	[4]				
RC	SEQUENCE OF 293-369 FROM N.A.				
RC	STRAIN=R6X;				
RC	MEDLINE=95020625; PubMed=7934910;				
RA	Pearce B.J., Yin Y.B., Masure H.R.;				
RT	"Genetic identification of exported proteins in Streptococcus				
RT	pneumoniae.";				
RL	Mol. Microbiol. 9:1037-1050(1993).				
CC	-I- FUNCTION: CELL WALL FORMATION.				
CC	-I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.				

```
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67873; CAA48073.1; -
DR EMBL: X67872; CAA48072.1; -
DR EMBL: M90527; AAA26956.1; -
DR EMBL: AE007349; AAK74536.1; -
DR PIR: S28038; S28038.
DR TIGR: SP0369; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme; Complete proteome.
FT ACCT_SITE 370 370 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT DOMAIN 658 683 SER-RICH.
FT VARIANT 124 124 T -> A (IN STRAIN R6).
FT VARIANT 386 386 V -> I (IN STRAIN 63915).
FT VARIANT 388 388 D -> E (IN STRAIN R6).
FT VARIANT 397 397 E -> K (IN STRAIN 63915).
FT VARIANT 523 523 M -> I (IN STRAIN 63915).
FT VARIANT 533 533 D -> I (IN STRAIN 45607).
FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).
FT VARIANT 657 657 N -> S (IN STRAINS 45607 AND R6).
SQ SEQUENCE 719 AA; 79758 MW; 5BD397E83B43BA6 CRC64;

Query Match 100.0%; Score 3484; DB 1; Length 719;
Best Local Similarity 100.0%; Pred. No. 2.5e-193;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIYDNKNOLIDLGSERVRNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFLRN 60
DB 54 KIYDNKNOLIDLGSERVRNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFLRN 113
OY 61 LOSNLSOGSSTLTQOLIKLTFYFSTSPOTSIRKAOEAMLAIOLEOKATKOEILTYINK 120
DB 114 LOSNLSOGSSTLTQOLIKLTFYFSTSPOTSIRKAOEAMLAIOLEOKATKOEILTYINK 173
OY 121 YMSNGNNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNQYDPSHPPEAODRRNLVL 180
DB 174 YMSNGNNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNQYDPSHPPEAODRRNLVL 233
OY 181 SEMKNOGYSIAEQYKAVNTPTDGLQSLKSASNPATMDNTLKVYNQVEETGYNLT 240
DB 234 SEMKNOGYSIAEQYKAVNTPTDGLQSLKSASNPATMDNTLKVYNQVEETGYNLT 293
OY 241 TGMDEVTVNDEQAKHLDIYNTDEYVAYPPDELOVASTYDVNSGKYIAOLGARRHOSSN 300
DB 294 TGMDEVTVNDEQAKHLDIYNTDEYVAYPPDELOVASTYDVNSGKYIAOLGARRHOSSN 353
OY 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEGVYDSTATIVHDEPNYPGNTPYVNMNR 360
DB 354 VSFGINQAVETNRDMSGTMKPTTDYAPALEGVYDSTATIVHDEPNYPGNTPYVNMNR 413
OY 361 GYFGNITITQYALQOOSRNPAYETLKNVGNRAKTFNLGLGIDYPSIHNSAISTTTSD 420
DB 414 GYFGNITITQYALQOOSRNPAYETLKNVGNRAKTFNLGLGIDYPSIHNSAISTTTSD 473
OY 421 KKYGASSEKMAAAYAPANGGYPYKTHKVVPSDSEKESPNYGTAMKETTAYMMTD 480
DB 474 KKYGASSEKMAAAYAPANGGYPYKTHKVVPSDSEKESPNYGTAMKETTAYMMTD 533
```

```
OY 481 MKKIVLTGTGTGNAYLAWLPQAGKGTGNTYDEIEIENHIKTSQVAPDELFAGYTRKYSM 540
DB 534 MKKIVLTGTGTGNAYLAWLPQAGKGTGNTYDEIEIENHIKTSQVAPDELFAGYTRKYSM 593
OY 541 AAWTGTGSRNLTPLVNGGLTVAAKVYRSMTYLSBESNPEDMWIPGGLVRNGEYFKKAR 600
DB 594 AAWTGTGSRNLTPLVNGGLTVAAKVYRSMTYLSBESNPEDMWIPGGLVRNGEYFKKAR 653
OY 601 STWNSPAPQAPPTSSSSSDSSTSSSTPTNNSTTTPNNNTPOQSNTPPOQON 660
DB 654 STWNSPAPQAPPTSSSSSDSSTSSSTPTNNSTTTPNNNTPOQSNTPPOQON 713
OY 661 PQAPQ 666
DB 714 PQAPQ 719

RESULT 2
PBPB_STROR STANDARD: PRT; 637 AA.
ID ID
AC 000573;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1A) (Fragment).
GN PONA.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92325042; PubMed=1624444;
RA Martin C., Biese T., Hakenbeck R.;
RT "Nucleotide sequences of genes encoding penicillin-binding proteins
RT from Streptococcus pneumoniae and Streptococcus oralis with high
RT homology to Escherichia coli penicillin-binding proteins 1a and 1b.";
RL J. Bacteriol. 174:4517-4523(1992).
CC -1- FUNCTION: CELL WALL FORMATION.
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M90528; AAA26958.1; -
DR PIR: B42893; B42893.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme.
FT ACCT_SITE 371 371 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT NON_TER 637 637
SQ SEQUENCE 637 AA; 70891 MW; A6D198BCEA603A63 CRC64;

Query Match 77.5%; Score 2698.5; DB 1; Length 637;
Best Local Similarity 87.5%; Pred. No. 3.5e-148;
Matches 511; Conservative 39; Mismatches 33; Indels 1; Gaps 1;

OY 1 KIYDNKNOLIDLGSERVRNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFLRN 60
DB 54 KIYDNKNOLIDLGSERVRNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFLRN 113
OY 61 LOSNLSOGSSTLTQOLIKLTFYFSTSPOTSIRKAOEAMLAIOLEOKATKOEILTYINK 119
```

```

114  LRGGGGLGASTLTQOLLKLTFFSTSTSDPTLSRKAQEMLAVOLEQKATKOEILTYIN 173
120  KYMNSNGVNGQTAONQNYGRDLNLSLPQLALLAGMPQAPQNYDPYSHPEAAQDRNLV 179
174  KYMNSNGVNGQTAONQNYGRDLNLSLPQLALLAGMPQAPQNYDPYSHPEAAQDRNLV 233
180  LSEMNQGYISAEQYERKAVNPTITDGLSLKSASNPVAMVDYKLEVINQVEEENGYNL 239
234  LSEMGQGYITAEQYERKAINPTITDGLSLKSANSYPPMDNYLKEVIDQVDEIGYNL 293
240  TTGMVDYTNVDOEAQRHMDIYNTEDEYVAYPPDELOVASTIYDVSNKGVIQAQARHSS 299
294  TTGMVEYTNVDSKVQORLMDIYNTEDEYVAYPPDELOVASTIYDVADGKVIQAQARHSS 353
300  NVSEGINQAVEINRDWGSIMKPTIYAPALEGYVDSTATYHDEPIYNTPTIYVND 359
334  NVSEGINQAVEINRDWGSIMKPTIYAPALEGYVDSTATYHDEPIYNTPTIYVND 413
360  RGYFNGITLOVALQOOSRNPVAVETLNKVLNRAKTFPLNGLGIDYPSIHYSNASSVTES 419
414  KSYFNGITLOVALQOOSRNPVAVETLNKVLNRAKTFPLNGLGIDYPSIHYSNASSVTES 473
420  DKRYGASSEKMAAAYAAFAFANGGTYYKPMYIHKVFSDSGESEKFSNVTGRAMETAYMT 479
474  DKRYGASSEKMAAAYAAFAFANGGTYYKPMYIHKVFSDSGESEKFSNVTGRAMETAYMT 533
480  DMKATVLYTGTRNAYLAMPQAGKTGTSNTYDELEHNIKTQYAPDELFRAGTRKKS 539
534  DMKATVLYTGTRNAYLAMPQAGKTGTSNTYDELEHNIKTQYAPDELFRAGTRKKS 593
540  MAVWTGYSNRFLPVLGNGLTVAAKYRSMATYLSGSDNEDNMI 583
594  MAVWTGYSNRFLPVLGNGLTVAAKYRSMATYLSGSDNEDNMI 637

```

```

RESULT 3
ID PPBA_BACSU STANDARD: PRT: 914 AA.
AC P39793:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1A/1B (PBP1) [includes: Penicillin-
insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan Tgase)];
DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
GN PONA.
OS Bacillus subtilis.
OC Bacteria: Firmicutes: Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxId=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 499-515.
RC STRAIN-168;
RX MEDLINE=95113769; PubMed=7814321;
RA Popham D.L., Setlow P.;
RT "Cloning, nucleotide sequence, and mutagenesis of the Bacillus
subtilis pona operon, which codes for penicillin-binding protein
(PBP) 1 and a PBP-related factor.";
RL J. Bacteriol. 177:326-335(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-168 / MARBURG;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
Serror P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
the sera and kds loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [3]
RP GROWTH REQUIREMENTS.
RC STRAIN-168;
RX MEDLINE=98389671; PubMed=9721295;

```

```

RA Murray T., Popham D.L., Setlow P.;
RT "Bacillus subtilis cells lacking penicillin-binding protein 1 require
increased levels of divalent cations for growth.";
RL J. Bacteriol. 180:4555-4563(1998).
RN [4]
RP SUBCELLULAR LOCATION.
RC STRAIN-168;
RX MEDLINE=99255546; PubMed=10322023;
RA Pedersen L.B., Angert E.R., Setlow P.;
RT "Septal localization of penicillin-binding protein 1 in Bacillus
subtilis.";
RL J. Bacteriol. 181:3201-3211(1999).
CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
(FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
SUBUNITS) (BY SIMILARITY).
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. PROBABLY FOUND ALL
OVER THE WHOLE CELL AT LOW CONCENTRATIONS. ALSO LOCALIZES TO THE
DIVISION SITE IN VEGETATIVE CELLS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS CONSTANT DURING GROWTH,
DECREASES DURING SPOULATION AND IS INDUCED APPROXIMATELY 15 MIN
INTO SPORE GERMINATION.
CC -1- PTM: THE PRODUCT EXPRESSED FROM THE TRANSLATION OF THE PONA GENE
APPEARS AS TWO BANDS ON A GEL (1A AND 1B), BUT THE SPECIFIC AMINO
ACID SEQUENCE OF EACH PROTEIN IS UNKNOWN.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- MISCELLANEOUS: CELLS LACKING THE PROTEIN REQUIRE INCREASED LEVELS
OF MG2+ OR CA2+ FOR GROWTH AND GERMINATION. APPROXIMATELY 50% OF
CELLS WITHOUT THE PROTEIN CONTAIN ABNORMAL FTZ2 RINGS, SUGGESTING
IT IS INVOLVED IN SEPTUM SYNTHESIS: INCREASED LEVELS OF MG2+ OR
CA2+ ONLY PARTIALLY ELIMINATE THE SEPTATION DEFECTS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
TRANSGLYCOSYLASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
TRANSEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U11883; AAA64947.1; -
DR EMBL: L47838; AAB38459.1; -
DR EMBL: Z99115; CAB14148.1; -
DR HSP: P02751; IFNA.
DR Subtilist; Bg10954; pona.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00041; fn3; 1.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR Pfam: PD001895; Transglycosyl; 1.
DR SMART: SM00060; FN3; 1.
DR Peptidoglycan synthesis: Cell wall; Transferase; Glycosyltransferase;
KW Hydrolyase; Multifunctional enzyme; Transmembrane; Signal-anchor;
KW Antibiotic resistance; Complete proteome.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 58 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT FT
FT DOMAIN 59 914 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 77 246 TRANSGLYCOSYLASE.
FT DOMAIN 329 662 TRANSPEPTIDASE.
FT ACT_SITE 390 390 ACYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 914 AA; 99562 MW; 6978E33DFE2423E6 CRC64;

```


Query Match 21.2%; Score 739; DB 1; Length 714;
Best Local Similarity 32.2%; Pred. No. 2,7e-35;
Matches 187; Conservative 119; Mismatches 237; Indels 38; Gaps 18;

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRIDITRILGAFLRN 60
DB 50 KIYDONGDEVASIYTEREPVSINEIPKQVREAFIATEDREFEHNHIDAKSVGRAVYRD 109
QY 61 -LQSNLSGGSTLTQOLIKLTFSTSDOTISRKQAMLAOLQEKAKKQELITYYN 119
DB 110 ILAGKVEGGTTTQOLAKIIFL---THDVTFLKTEVITAINLERDYSKDLLEYLN 166
QY 120 KYVMSNGYCMQTAONYKDKLNLNLSLPOLALLAGMPAPNOYDPSPHEAADRNLV 179
DB 167 QLEFGHGVYIGQASHYIFPKKEVKDLTVSGAVLAIAIPKAPSTYSPILHDKKERDIT 226
QY 180 LSEKNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNLYKEVINQVEE--TGY 236
DB 227 LGMNDQGYISAKEAVTAQGR--TLGLH-VKQSETP-WPDSYIDLVIKAEEDKYSISGE 282
QY 237 NLTGMDVYTNVDOEAKHLMIDYNTDEVAAFPDDELQVASTIVDYSNCKVIAOLGARH 296
DB 283 QLLQGGTTIKVPLDSKLOKTAQVMEKGSYYPETDQNAESGSAVFNKKTGVEALIGRD 342
QY 297 OSSNVSFGINQAVETNRDMGSTMKPTIDYAPALEGYVDSTATIVHDEPNYPGTNPVY 356
DB 343 YTSK---GYNR-VTAVROPOSTFKPLAVYGPMAQEKKF-KPYSILKDELQSY-GDITP-K 395
QY 357 NMDRGYFNTTIOYALQOSHNPAVETLNKVGILNRAKFTNGLDIDPSPHYSNAISSNT 416
DB 396 NYDSREGEVYTMADATYTSKNAPAVWTLNEIGVETGSKYLKANGCIDIPDEGLALALGG- 453
QY 417 TESDKKYGASSEKMAAAYAFANGCTYKPMYTHKVVFSGSGE-KEPSNVTGAMKETTA 475
DB 454 ---LEKGVSPLOLAGAFHTFAANGTYTEPFPTSSIDDEGETIADHKEGKRVFSKOTS 509
QY 476 YWMTDMKVTLYTGTGRN-AYLAWLPQAGKTGTSNTDEIENHIKTSQVAPDELAGY 534
DB 510 WNNTRMLQOVYKKGTISGTYHGDLL-AGTGTSTYTG--VSGATK-----DAMFAGY 558
QY 535 TRKYSMAVWTGYSNR-LTPLVNGLTVAAKVYRSMTYLSGE 574
DB 559 TPKITGAVVMGVDKTDQNHYLKAGSSYPTRLFKDILLQAGE 599

RESULT 5
BPBA_AQDAE
ID BPBA_AQDAE STANDARD; PRT; 726 AA.
AC 066874;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Penicillin-binding protein 1A (PBP-1a) (BPPIa) [Includes: Penicillin-
insensitive transglycosylase (PC 2.4.2.-) (Peptidoglycan TGase);
DE Penicillin-sensitive transpeptidase (PC 3.4.4.-) (DD-transpeptidase)].
GN MRCA OR POMA OR AQ_624.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5.
RX MEDLINE=98196666; PubMed=9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Tenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus".
RU Nature 392:353-358(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: AE000699; AAC06835.1; -
DB InterPro: IPR001264; Transglycosyl.
DB InterPro: IPR001460; Transpeptidase.
DB Pfam: PF009912; Transglycosyl. 1.
DB Pfam: PF009905; Transpeptidase; 1.
DB ProDom: PD001895; Transglycosyl. 1.
DB Peptidoglycan synthetase; Cell wall; transferase; Glycosyltransferase;
KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
KW Signal-anchor; Antibiotic resistance; Complete proteome.
FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 4 24 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 25 726 PERIPLASMIC (POTENTIAL).
FT DOMAIN 45 213 TRANSGLYCOSYLASE.
FT DOMAIN 379 662 TRANSPEPTIDASE.
FT ACT_SITE 432 432 ACYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 726 AA; 81824 MW; 37f756397c9d7b38 CRC64;

Query Match 20.4%; Score 712; DB 1; Length 726;
Best Local Similarity 30.7%; Pred. No. 1e-33;
Matches 207; Conservative 109; Mismatches 227; Indels 132; Gaps 21;

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRIDITRILGAFLRN 60
DB 46 EYVDARKRLVGTIGIQKRFVYSIDKIPYEHVINFVAREDRNFMHFQIDPAIVRAIALVN 105
QY 61 LQSNLS-OGSSTLTQOLIKLTFSTSDOTISRKQAMLAOLQEKAKKQELITYYN 119
DB 106 YRGRIVQGGSTTTQOLAKMFL---TRERTLEKRIKIALLAIKIERTFPKKKIMLEYLN 162
QY 120 KYVMSNGYCMQTAONYKDKLNLNLSLPOLALLAGMPAPNOYDPSPHEAADRNLV 179
DB 163 QYLGSGAYVEAAQVYFGHWHWELSLDEAALALAPKAPAKNPYHBERALQDRNLV 222
QY 180 LSEKNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNLYKEVINQVEETGYNL 239
DB 223 LKRRLEGGYITPQGEAAVAKPLT---VKREKKY-KFGDYFLDMVKSYVFNKYGELAY 276
QY 240 TTGMDVYTNVDOE---AQHLMND----- 259
DB 277 KGRILKITYTIDLYQKIAQKSLSEGILKRVAKIIGLPPLPKSEEDMELAYEKEADQLRLKR 336
QY 260 --TY-----NTD-----BYVAV-----PDDELQY 276
DB 337 GKTYVAKILYDGNFMKVEIHGKKLKEIGLNTGKHVYFVKLGGNRAELIIPDLGSL 396
QY 277 ASTIVDYSNCKVIAOLGARHQSSNVSFGINQAVETNRDMGSTMKPTIDYAPALEGYVDS 336
DB 397 VS-IDVKTEIKAIIVGR---SYASQFNRKAVKALNPOPSAIIKPYI-YLSALLKGMTQI 450
QY 337 TATIVHDEPYNYP--GTNPVYKMDRGYFGNITLOVALQOSRNPAVETLNKVGILNRAKT 394
DB 451 STIDASSKPYDPSKGEDMIPKNYDEKEYGNVTLRYALASHIMTAVALNLDKGFGFELYLE 510
QY 395 FLNGLGIDYPSIHYSNASSNTTESDKKYGASSKMAAAYAFANGSTYKPMYIHVVF 454
DB 511 VGKKVGLDNLKPYSTSLV-GTVE-----VTPQLTAAOVFANLGTCECKPFIKKIYD 562
QY 455 SDSEKEFFSNVG--TRAKKETTAVMYDMKVTLYTGTGNAYLAWLPQAGKTGTSNYTD 512
DB 563 ENGEVYLE-ENVPCEBEVLPRPETHRVYDMLRAVYLBETARASVLDIIVAGKGT---FD 618
QY 513 EEIENHIKTSQFAVAPDELAFAGYTRKYSMAVVTGYSNRLPLVVG---NGLTVAAKVRSSM 569
DB 619 D-----FODAMFVGFSPYIVTGVWGVGVAKS--LGKHMGSRYALPPIWIDM 664
QY 570 TYLSEGSNPDWNP 584

DB 665 KVVTRMYPNEDEFLP 679

RESULT 6

PPPA_NEIFL STANDARD; PRT; 798 AA.

ID PPBA_NEIFL 087626;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [Includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan TGase)]; MRCA OR PONA.

GN Neisseria flavescens.

OS Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

OX NCBI_TaxID=484;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-NRL 30009;

RA Ropp P.A., Nicholas R.A.;

RT "Nucleotide sequence of the ponA gene encoding penicillin-binding protein 1 of Neisseria flavescens."

RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE SUBUNITS) (BY SIMILARITY).

CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.

CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By similarity).

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE TRANSGLYCOSYLASE FAMILY.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE TRANSPEPTIDASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Its usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF087677; AAC35856.1; .

DR InterPro: IPR001264; Transglycosyl.

DR InterPro: IPR001460; Transpeptidase.

DR Pfam: PF009012; Transglycosyl; 1.

DR Pfam: PF009005; Transpeptidase; 1.

DR Prodom: PD001895; Transglycosyl; 1.

KM Peptidoglycan synthesis; Cell wall; Transferrase; Glycosyltransferase; KM Hydrolyase; Multifunctional enzyme; Transmembrane; Inner membrane; KM Signal-anchor; Antibiotic resistance.

FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN).

FT FT (POTENTIAL).

FT DOMAIN 31 798 PERIPLASMIC (POTENTIAL).

FT DOMAIN 50 218 TRANSGLYCOSYLASE.

FT DOMAIN 413 699 TRANSPEPTIDASE.

FT ACT_SITE 460 460 ACYLATED BY PENICILLIN (BY SIMILARITY).

SO SEQUENCE 798 AA; 87703 MW; 0DDCD6FD5953AA CRC64;

Query Match 19.5%; Score 680.5; DB 1; Length 798;

Best Local Similarity 25.8%; Pred. No. 7.5e-32;

Matches 203; Conservative 129; Mismatches 277; Indels 179; Gaps 20;

OY 2 IYDNKNOLADIQSERVNAQNDIPTDLVKAIVSIEDHREFDHCIGTIRILGAFRLN 61

DB 52 IYSSGQVIGVYGEQRREFTKIDPEKILKDAVIAEDKRFYDHMGVDWGVARAVIGNV 111

OY 62 OSNSLD-QGSTLTLOQLIKLTVSTSTSDQTSKRAQEAWLAIQLEOKATKOELLTYINK 120

DB 112 MAGGVQSGASTITQYAKNPFYLS---SERSFTFKFEALLATYIEBSLKDKLTLEYFNG 168

OY 121 VYWSNGNYGMOTPAONYGYKDLNNLSLPOLALLAGMPQAPNOYDPSHPBAODRNLV 180

DB 169 IYLGQRAYGFASAAQYTFNKNVNDLTLEAAMLAGLPKAPSAVNPYVNEBRAKLRQAYLL 228

OY 181 SEKKNGCYISAEYEAQVMTPTITDGLQSLKASNYAVYNDN-----YKKEYNO-VVEE 233

DB 229 NNMLEGKMTTLQORD-----QALKELHYERFVQNDIQSALVYAEKARQELFER 277

OY 234 TGYNLITGMDVT-----NDQEA 253

DB 278 YGEDATYQGFKYTTTVDTAHQVATEALKVLRNPDGSSRYGAENYIDLSKSDVEEIV 337

OY 254 QKHLDIYNTDEVY-----AYPDEL 274

DB 338 SOYLSTLYTVDKMIPAVVLEASRKGYQIOLPSGRKYTLNNHALGFARAANNKMGDDRI 397

OY 275 QVAST-----VDVSNKGVIAQLGARHQSNVSGINQAVET 311

DB 398 RRGSVIRVKGSGDTFTVVOEPLQALVSLDAKTGVALVGGYDYHSKT---FNRATQA 454

OY 312 NRDMGSTMPTIDYAPALEGYVDSTATVHDEPNYPGTNPVYVW-----DRGYFGNI 366

DB 455 MQPGSTFPRFL-YSAALAKGM--TASTMINAPISLPKANGKAMNKNKSDGRAGVI 511

OY 367 TLQYALQSRNPVAVETLNKVGILNRAKTLFNLGIDYPSIHYSNMISSTTESDKKYGAS 426

DB 512 TLROALTASKNWSIRILMSIGVYAOQYTORGFKEIPASLSAALAGET-----T 565

OY 427 SEKMAAAYAPANGCTYYPYTHKVFPSDSEKE-----ESNVGTRAMKETTYAMMD 480

DB 566 PLRIAGISVFANGYKYSAAHYIDKYDSQGRIRAOQPLVAGENAPQIDPRANATIMK 625

OY 481 MKTVELTYGTGRN-AVLAMLPQAGKTGTSNYPDEIEHNIKTSQFAPDELPAQYTRKYS 539

DB 626 INQDVVRVTAAGATNLGSDIAGKTGTN-----DN-----KAMVGVGFNPV 670

OY 540 MAVMTGYSN-RLTPLYNGCLTVAAKYRSMYTLSEGSNPEDMNIPDEGLRNGEYFKNG 598

DB 671 TAVYIGFDPRSRMGRCAGYGTIAVPWVEYIGFALKGTGVPKAPEGVVTNGEYV--- 727

OY 599 AASTWSPAPQPPSPRESSSSDSTSSQSSSTPSTNNSTTTPNNTNQOSTPDDON 658

DB 728 -----MKRMTTSSDLALDNGIRPRPTQPA--RRAVPENRRRAESNTAPAREE 775

OY 659 QNPQAP 666

DB 776 SDFTPLP 783

RESULT 7

PPPA_PSEAE STANDARD; PRT; 822 AA.

ID PPBA_PSEAE 007806;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [Includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan TGase)]; MRCA OR PONA OR PA5045.

GN Pseudomonas aeruginosa.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OX NCBI_TaxID=287;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 15692 / PA01;

Handfield J., Gagnon L., Dargis M., Huletsky A.;

[illegible]

QY	61	LOSNSLO--GGSTLTQOQLKLTLEFSTSTSDQFISSKRAQEAAMLAIOLEOKAKOELLTYTIN	112
Db	109	LKSGHIDGGSTLTITMOQAVK-NYF--LTNERSESKINIEILLALIOEQLTRKDELLEYLVN	165
QY	120	KVYMSNGYNGQOTAAONYVYGKDLNNLSIPOLLALLAGMPOANQOYDPXSHPEAADRRNLV	179
Db	166	KIYLGNRAHYGIEAAAOYYTGGFPINDLSIAEAMAMLAGPKPASKRNPLVNTPTSTERRNMI	225
QY	180	LSEMKNOGYISAEQYKAVNPTITDGLQSLKSNASNPAYMDNTLKEYINOVEEET---GY	236
Db	226	LEBMUKIGEIFDQGRQAAVEEPINASYHVQTPELNABYIAEMARAEMVGRGSEATYEGY	285
QY	237	NLLTTGMDVYTN-----VQGEAQ-----KHL-----	257
Db	286	KVITTVRSDDLQMAASQSVRDGLIDYDQRHGRGPETRLPGQTRDAMLKHHGQOORSIGLE	345
QY	288	-----WD-----IYNDEVAAPDDELOY	276
Db	346	PAIYTOVEKSGIMVWTRDCKEAEATWMSKKARFPLSNSKGPMPROPADVAQAGDIDRY	405
QY	277	-----ASTIVDVNSGNYIAQGA-RHOSSNVSEGINOAIVETNRDWS	317
Db	406	QROEGCTLRFVQIDPAQAGSALISLDPKQGAISLIVGGFSEGSN---YNAIDAKKRQGS	461
QY	318	TMKPTTIDYAPALEGYVDSTATIYHDEPYNT-----PGTNPVYNNMDRGTYEGN	365
Db	462	SEKPEI-YSAALDNGF--TAASLVNDAPIVFVDEYLDKVMRPKNDTNT-----FLGP	510
QY	366	ITLQVALOOSNNVPAVETLNVKGNRAKTELTNGIDIDYPSI--HYSNAISNTTESDKY	423
Db	511	IPLRALKTKSNMWSIRVLOGLGIERAISTYTKGFQRODELPRNFSLALGTAT-----	563
QY	424	GASSEKMAAAAYAAPANGGTYKKPMYIHKVVFSDS-----SEKEFSN	464
Db	564	-VTPEIENGAMSVFANGSKVKNPVIERIESRDOQVLYQANPRPVVEQVADAEADAGN	622
QY	465	VG-----TRAKETTVAMTDMMKTYVLYTGGR	492
Db	623	PGDDEHPESAEGESIEAQVAARQOTFEPTPAERILIDRTAGTAVYIKMSLMQDVIKRGTR	682
QY	493	NAY-TAMLPQAKGTGTSYNTDELENIHKTSQFAPAPDELFAGYRKYKSMVWGYSNRLT	551
Db	663	RALAKKRTDLGKGTGN-----DSKGGWFSGVNSVDVTYVWVGFPDQPET	727
QY	552	PLVGN---GLTFAAKVYSMTYISSEGSNPEEDWNIPEGILRNGEFGFYKNGARSWTNSPAP	608
Db	728	-LGRREGGIVALPIWIRYRGFALKKPKPHITMAPEPGT-----VSLRIDPVTGKSAAP	779
QY	609	QQPPSTESSSSSDSSTSSQSSSTTPSTIN	636
Db	780	GTPGAYFEMFKNED-----TPPSVN	799
RESULT 8			
PBPA_NEILTA STANDARD; PRT; 798 AA.			
ID	16	16-OCT-2001 (Rel. 40, Created)	
AC	087579;		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Penicillin-binding protein 1A (pbp-1a) (pbpla) [includes: penicillin-in-		
DE	insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan Tcase)]		
DE	Penicillin-sensitive transpeptidase (EC 3.4.-.-) (Dd-transpeptidase)]		
GN	MCR4 OR PONA.		
OS	Neisseria lactamica.		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=486;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NRL 3716;		
RA	Ropp P.A., Nicholas R.A.;		
RT	"Nucleotide sequence of the ponA gene encoding penicillin-binding protein 1 from Neisseria lactamica.";		

```

RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY).
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (by
CC similarity).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSEPTIDASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF085689; AAC35363.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl. 1.
KW Peptidoglycan synthetase; Cell wall; Transferrase; Glycosyltransferase;
KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
KW Signal-anchor; Antibiotic resistance.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT FT 31 798 (POTENTIAL). (POTENTIAL).
FT DOMAIN 50 218 PERIPLASMIC (POTENTIAL).
FT DOMAIN 378 700 TRANSGLYCOSYLASE.
FT ACI_STATE 461 461 ACYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 798 AA; 88108 MW; 389C672B86935D6 CRC64;

```

```

Query Match 18.4%; Score 640; DB 1; Length 798;
Best Local Similarity 26.0%; Pred. No. 1.6e-29;
Matches 203; Conservative 126; Mismatches 265; Indels 186; Gaps 24;

```

```

OY 2 IYDKKNQIADLGSERRNAQANDIPDLVKAISIEDHREFDRGIDITIFILGFLRNL 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 52 VYSADGKIIGYGEORREFTKIGDPEVLNAVAEDKRFYQHMGVDVWGVARAIVGNI 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 62 QSNELQ-GGSLTQOLIKLYFTSTSDQTSRKAOEAMLAIOLEOKATKOEILTYTINK 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 112 VSGSMQASISITTOQVANKFLS---SEKTFKRFNEALLKIKIQSISKIKILELYNQ 168
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 121 VYKNGNGMOTAAQNYKQDLNLSLPQALLAGMPAPQNYDPYSHPAQAODRNLVL 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169 IYLQRAVGFASAOIYFNKDVRLTLAEAMLAGLPAKPAVNPYVPERAKILQKYTL 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 181 SEMKNGCISAEQYKAVNPTITDGLSKASNPAYMDN-----YLKEVING-VSEE 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 229 NNMLEEKYITLQORQALNEEL-----HYERFQKIDQSALVAEMVRELXEK 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 234 TGYMLLTGMDVYTNV-----DOEAQ--- 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 278 YGEDAVYQGLKLYTTRVDHOKATEALRKALRNDRSSYRGAESYIDLGRDEAEAV 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 255 -KHLMDIYNDVEY-----AYPDE----- 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 338 SOYLSGLYTVDKMVPAAVLDTYKKKNVLIQPGCKRVTLDRRALCFARAVDNEMKEDR 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 274 -----LQVASTIYVSNKVIADLGARHQSNNSPFGINQAVE 310
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 398 IRRGAVIRNRNGRMAYVQEPPLQGLVLSIDAKTGAVRALVGGYDFHSKT---FNRAIQ 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

OY 311 TNRDWSMTKEITDYAPALEXYVDSTATIVHDEPNYPGNT-----TPVYNDRGYFG 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 455 AMRPGSTFKFPEV-YSAALSQGM--TASTMINDAPISLPGKPGNSWTP-KNSGRYSG 510
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 365 NITLQALQOSRNPFAVETLTKVGLNKRKFTLNLGIDIPYPIHNSALSSMTSDSKYK 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 511 YITLQALTLASKNNVSIILMSIGVIAQYIRRFGRFSELPVLSALTLGT----- 565
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 425 ASSEKMAAAYAFANGCTYYRPMYTHKVFSDGSEKE-----PSNVTFRAMKETAYAM 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 566 -TPRLIAAYSVFANGGRVSSVYIDKTYDESGRLRAQMPQVLAQGNAPQAIIDPRNATIM 624
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 479 TDMKVTLYTGTGRNA-YLAWLPOAGKTGTSNYDEEIEHNIKTSQVAPADELEPAGYTRK 537
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 625 YKIMQDVYRGVGTARGASALGRSDIAGKTGTNT-----DN-----KDAWFGFNPD 669
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 538 YSMAMVTCYSN-RLTPLVGNGLTYAAKYRSMNTYLSGSPNEDMNIPEGLYR-NGEEVF 595
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 670 VVTAVYIGDFDKPSMGRAGYGTIAVPWVDYMRPALKGRPGKMKMPDGVVAGNGEYIM 729
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 596 KN-----GARSTWNSPAPQGPS-----TESSSSSDSSTSSQSSSTPSTNNST 639
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 730 KEHWYTDGLMLDNGGAAPQ--PSRRVKEDDGAAGGCGRADESDRQDMQETVLPSTNT 787
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 9
PBPA_NEIMA STANDARD; PRT; 798 AA.
ID PBPA_NEIMA
AC 005194;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [includes: Penicillin-
DE insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan tase);
DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (Dd-transpeptidase)].
GN MRCA OR PONA OR NMA0653 OR NMB1807.
OS Neisseria meningitidis (serogroup A), and
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699, 491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=97252514; PubMed=9098083;
RA Ropp P.A., Nicholas R.A.;
RT "Cloning and characterization of the ponA gene encoding penicillin-
RT binding protein 1 from Neisseria gonorrhoeae and Neisseria
RT meningitidis."
RL J. Bacteriol. 179:2783-2787(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holtroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrett B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tectelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Hat D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Clecio A., Parksey D.S., Blair E., Ciltone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,

```


Query Match 18.2%; Score 633; DB 1; Length 798;
 Best Local Similarity 26.3%; Pred. No. 4.1e-29;
 Matches 191; Conservative 115; Mismatches 248; Indels 166; Gaps 19;

2 IYDNKNOLIADGSEFRVNAQANDIPTDVLKAIYSIEDHREFDHRGIDITRIILGAFRLN 61
 DB 52 IYSADGEVIGIEGRREPRFKIGDPEVILRNAYIAEDKRFQHMGCVDWYARAVNGMT 111

62 QSNLSIQ-GGSLTLOOLIKLYEFSTSDQTSRKAQEAWLAIOLBKATKOEILTYIYNK 120
 DB 112 VAGVGSGASTITQOVAKNFYLS---SEKTFTRKFNEALLAVKIQSLSKDKILLEYFNQ 168

121 VYMSGNGMGTAAQONYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAAQDRRLNVL 180
 DB 169 IYLGRAAGFASAAQIYFNKDVRELTLEAVAMLAGLPRKAPSNAPNIVPERAKLRQKYL 228

181 SEMKNGYISAEQYKAVNTPTDGLQSLKSASNPAYMDN-----YLKEYINQ-VEEE 233
 DB 229 NNMLEEKMTITLOQRDQALNEEL-----HIERVQKIDQSALYAEWNRQELYEK 277

234 TGYNLLTGMVYTNVDOEAK----- 255

278 YGEDAYTQGFKYTTVRDHDQKVAATEALRKALRNFDROSSYRGASYSIDLKSGDEVEETV 337

256 --HLMDIYNTDE----- 265

338 SOYLEGLTYTVDKMPAIVLDYTKRRNVYIOLPSGKRVTLIDGSLGLFAARAANNKMGESR 397

266 -----YVAVPDEELQVASTIVNSNGKYIAQAGARHOSNSVFGINQANE 310

388 IRRGSVIRNRNGGKMWVVOEPLLOATVLSLDAKTGAVALGVDGFHSKT---FNRAAQ 454

311 TNRDMSGTMKPTTDYAPALEYGVYDSTATTIVDEPYNYP--GTNTPVY---NMDRGYEFGN 365

455 AMRQGSSTFKPEPI-YSAALSCKM--TASTMVDAPISLPGKANSQVMPKPSKSDGRYSY 511

366 ITLOALDOORNPVAVETLNKVLNRAKTFELNGLIDIPYSIHSNAISSNTTESDKKYGA 425

512 IYLRQALASKNMVSIIRILMSIGVYAHGYIORFGFKSELPASLSMALGTGET----- 565

426 SSEKMAAAYAFANGCTGYKKPMYIHKVYFSSQSEKE-----FSNVGTRAMKETTAYMMF 479

566 TPLKTAIEIYVFANGGYRVSSHVIDKIGSDGLRAQOMPLVAGONAPADIPRANAYIMY 625

480 DMKMTVLTLYGTGRN-AYLAMLPOACKGTGTSNTDEIENHIKTSQFVAPDELFAGYTRKY 538

626 KIMQDVAVVGTRAGAAALGRSDIAGKTGTN-----DN-----KDAWFGVFNPDV 670

539 SKAVVTGTSN-RLPPLVGNGLTVAAKYRSMNTYLSGGSNPEDMNIPEGLY-RNGEYVFK 596

671 VTAAVYIGDPKPSKMGKAGGFTIAVPVWYMRFAKLGQCKGMKVPGEVSSNGEYYMK 730

RESULT 12
 PBPA_VIBCH STANDARD; PRT; 825 AA.
 ID PBPA_VIBCH
 AC O9KN05;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Penicillin-binding protein 1A (PBP-1a) (PBPA) [includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase); Penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
 GN MRCA OR PONA OR VC2635.
 OS Vibrio cholerae.
 OC Bacteria, Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN SEQUENCE FROM N.A.
 RP STRAIN=EL TOR N16961 / SEROTYPE O1;
 RC MEDLINE=20406633; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*.";
 RL Nature 406:477-483(2000).
 CC -I- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY).
 CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
 CC similarity).
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSPEPTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as their content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE004330; AAF5776.1; ALT_INIT.
 CC DR TIGR: VC2635;
 CC DR InterPro: IPR001264; Transglycosyl.
 CC DR Pfam: PF00912; Transglycosyl; 1.
 CC DR Prodom: PD001895; Transglycosyl; 1.
 CC KW Peptidoglycan synthetase; Cell wall; Transferase; Glycosyltransferase;
 CC KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
 CC KW Signal-anchor; Antibiotic resistance; Complete proteome.
 CC FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 7 27 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
 CC FT FT (POTENTIAL).
 CC FT DOMAIN 28 825 PERIPLASMIC (POTENTIAL).
 CC FT DOMAIN 48 216 TRANSGLYCOSYLASE.
 CC FT DOMAIN 413 752 TRANSPEPTIDASE.
 CC FT ACT SITE 471 471 ACYLATED BY PENICILLIN (BY SIMILARITY).
 CC SQ SEQUENCE 825 AA; 91986 MW; 8A953DA2650B9D2 CRC64;

Query Match 17.7%; Score 617; DB 1; Length 825;
 Best Local Similarity 24.3%; Pred. No. 3.6e-28;
 Matches 200; Conservative 128; Mismatches 220; Indels 276; Gaps 26;

1 KIYDNKNOLIADGSEFRVNAQANDIPTDVLKAIYSIEDHREFDHRGIDITRIILGAFRLN 60
 DB 49 QYFSDQGLIAQFGRKRRIPLKLEMPKRLIAVATIEDSRVYEHGFGPIITRAFAV 108

61 LOSNSL-OGGSTLTOOLIKLYEFSTSDQTSRKAQEAWLAIOLBKATKOEILTYIYN 119

109 IASGASQASGASTITQOLANFNFLS---NEKKYWRKKEIFITIHIEQLSKOEILTYIYN 165

120 KYVMSGNGMGTAAQONYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAAQDRRLNVL 179

166 KIYLGIRSVGGAQAQAVYRKEVKDULTEIGALINGLPRAPSTMPPIYSVEBATNRNRY 225

180 LSEKNGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEV-----INOVEEET 234

226 LQRMDEKTYITKAEYDAARAEPVLPKRYHGAETELNAP-----YVAEIRAMVYRGEBA 280

235 GYNLLTGMVYTNVDOE-----AOKHLMDI----- 260

281 AY---TSGMNVITYVDSKLQRAAQAALNNLLAYDERHRYRGAERKELMVOGNPANSSTQL 337

ID PBPB_HAEIN STANDARD; PRT: 853 AA.
AC P31776;
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 40, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) (Penicillin-binding protein A) (includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan Gase); Penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)).
GN MCA OR PMA OR H10440.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=92009183; PubMed=1916268;
RA Tomb J.-F., El-Hajj H., Smith H.O.;
RT "Nucleotide sequence of a cluster of genes involved in the transformation of Haemophilus influenzae Rd.";
RL Gene 104:1-10(1991).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M., Uettermann J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Weidman T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
RL Science 269:496-512(1995).
RN
RP CHARACTERIZATION.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=96074300; PubMed=7592463;
RA Sharma U.K., Dwarkanath P., Banerjee N., Town C., Balganesi T.S.;
RT "Expression and characterization of the ponA (ORF 1) gene of Haemophilus influenzae: functional complementation in a heterologous system.";
RL J. Bacteriol. 177:6745-6750(1995).
RN
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B., Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-429(2000).
CC
CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE SUBUNITS) (BY SIMILARITY).
CC
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By similarity).
CC
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE TRANSGLYCOSYLASE FAMILY.
CC
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE TRANSPEPTIDASE FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M62809; AAA25007.1; -;
DR EMBL: U32727; AAC22099.1; -;
DR PIR: JH0438; JH0438.
DR TIGR: H10440; -;
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Cell wall; Transferrase; glycosyltransferase;
KW Hydroxylase; Multifunctional enzyme; Transmembrane; Inner membrane;
KW Signal-anchor; Antibiotic resistance; Complete proteome.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 7 27 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT
FT DOMAIN 28 853 PERIPLASMIC (POTENTIAL).
FT DOMAIN 37 205 TRANSGLYCOSYLASE.
FT DOMAIN 387 681 TRANSPEPTIDASE.
FT ACT_SITE 441 441 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT CONFLICT 33 33 L -> LPSVERTLKTVE (IN REF. 1).
SQ SEQUENCE 853 AA; 94221 MW; FC0846096CDB663B CRC64;

Query Match 17.0%; Score 591.5; DB 1; Length 853;
Best Local Similarity 29.2%; Pred. No. 1,le-26;
Matches 166; Conservative 99; Mismatches 176; Indels 127; Gaps 20;

QY 1 KIYDKKNIQDLIGSERVNAQNDIPDLYKAIYSIEDHREPHRGIDTRILGA-FLR 59
DB 38 QITRQGLKEGEGQKRIPIVKLADVORLIDAFATEDSRFDHGGIDPGLIAALFYA 97
QY 60 NLOSNLGGSTLTQOLIKLTYFTSTSDQTSRKAQAMLAIOLEOKATNOEILTYIN 119
DB 98 VSNCGASQSGASTITQOLARNFLL--TSEKTIIRKAREAVLAVEIENTLNKOEILEYLN 154
QY 120 KYVNSNGYGMQTAQNTYGGDLNLSLPOLALAGMQAPNQDYPYSHPEAODRMLY 179
DB 155 KIFLYRSYGVAAAOYQTFEGLSLNLTLEMAIITGLPKASTMNPYLSKSEERRNV 214
QY 180 LSEKMNQYISAEOYERKVNPIPTDGLQSLKSASVPAVMQNYLKEVINQVEEFGYML 239
DB 215 LSRMLDEKTYISKEEDALKEPIYASYGAKFRPADYVTEVROEMKRRGEENAY-- 271
QY 240 TTGMQVYTNV--DQ-EAOK-----HLMD-----LYNTDEVAY-- 269
DB 272 TSGYVFETTVLSKQDAEQAKAVRNRLIDYDMRGYRGAPLMOKNKAAAMDRIYGLR 331
QY 270 -PDDELQVASTVDVSNQ--KVIAQLGAR-----HOSSN- 300
DB 332 LPDSEPTIPAAVITGIVKGADILLASGEKMTLSTNAMRWTKRSNPKYKGEQIWHQRANG 391
QY 301 -----VSFG-----INQAVETNRDMSGTKPTITVDAP 327
DB 392 EMQLQIPANSALVSLNSDNGALFVAVGGEYSQSKFNRAQSLVQVSSSTKPTI-FAA 450
QY 328 ALEVGYVSTATTIVHDEPNY--PG-----TNPFVYMDRQFGNITLQVALQDSRNP 379
DB 451 ALERKL--TLSSVLQDPSITISQKPGQKMMQPNSP--DR-YDSRMLRLRGLGSKNII 503
QY 360 AVETLNKYLGNRAKTFPLNGLGIDIPYSIHYSNAISSNTTESDKKYGASS--EKAAYAA 436
DB 504 AIRAIGTIGIDFTEFFLQRFGRKRDQYFASALIA-----LGAASFTPLEMARAYAV 554
QY 437 FANGCTYKPMYIHKVVPDSSEKFEFN 464
DB 555 FDNQGFLEPIYIEIKIQDNTGKOLFIA 582

RESULT 15
PBPB_VIRCH

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 04:24:19 ; Search time 86.16 seconds

(without alignments)
1337.217 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIYDNKNQIADLGSERRV.....TQOSNTTPDOQNQNPQAP 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.todent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.rvirus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	719	2	Q9REU0 streptococc
2	3468	99.5	719	2	Q9RET9 streptococc
3	3403	97.7	719	2	Q9RET9 streptococc
4	3329	95.6	719	2	Q9WVW0 streptococc
5	3314	95.1	719	2	Q9RET8 streptococc
6	3299	94.7	719	2	Q9WVW1 streptococc
7	3274	94.0	719	2	Q9RET4 streptococc
8	3247	93.2	719	2	Q9RET6 streptococc
9	3246	93.2	719	2	Q9RET6 streptococc
10	3245	93.1	719	2	Q9RET6 streptococc
11	3243	93.1	719	2	Q9RET6 streptococc
12	3241	93.0	719	2	Q9RET6 streptococc
13	3231	92.7	719	2	Q9RET5 streptococc
14	3192	91.6	719	2	Q9RET5 streptococc
15	3134	90.0	718	2	Q9RET7 streptococc

17	2820	80.9	608	2	054951 streptococc
18	2791	80.1	728	2	09F2G7 streptococc
19	2696	77.4	727	2	09F2G4 streptococc
20	2432	69.8	527	2	070037 streptococc
21	2091	60.0	398	2	Q9R7M6 streptococc
22	2088	59.9	398	2	087102 streptococc
23	2037.5	58.5	721	16	099YV1 streptococc
24	2029	58.2	398	2	Q9R7M5 streptococc
25	2015	57.8	398	2	087108 streptococc
26	2011	57.7	398	2	087105 streptococc
27	1936	55.6	398	2	Q9R7M7 streptococc
28	1926	55.3	398	2	087103 streptococc
29	1921	55.1	398	2	Q9R7M8 streptococc
30	1911	54.9	793	2	Q9EXM9 streptococc
31	1881	54.0	398	2	087107 streptococc
32	1876	53.8	398	2	087104 streptococc
33	1757	50.4	400	2	087106 streptococc
34	1751	50.3	778	2	Q9EXN1 streptococc
35	1657.5	47.6	664	16	Q9C123 streptococc
36	1592	45.7	310	2	052741 streptococc
37	1583	45.4	310	2	052743 streptococc
38	1573	45.1	310	2	052739 streptococc
39	1571	45.1	310	2	052737 streptococc
40	1567	45.0	310	2	052738 streptococc
41	1564	44.9	310	2	052740 streptococc
42	1558	44.7	310	2	052742 streptococc
43	1454	41.7	310	2	Q93N72 streptococc
44	1449	41.6	310	2	052744 streptococc
45	1201	34.5	826	16	Q92AB6 streptococc

ALIGNMENTS

RESULT	ID	Q9REU0	PRELIMINARY:	PRT:	719 AA.
AC	Q9REU0	01-MAY-2000 (TREMblrel. 13, Created)			
DT	01-MAY-2000 (TREMblrel. 13, Last sequence update)				
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)				
DE	PENICILLIN-BINDING PROTEIN 1A.				
GN	BPPIA.				
OS	Streptococcus pneumoniae.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1313;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SP 1261;				
RX	MEDLINE=21432820; PubMed=11549185;				
RA	Ferroni A., Berche P.;				
RT	"Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children."				
RL	J. Med. Microbiol. 50:828-832(2001).				
DR	EMBL; AF210745; AAF17255.1;				
DR	InterPro; IPR001264; Transglycosyl.				
DR	InterPro; IPR001460; Transpeptidase.				
DR	Pfam; PF00912; Transglycosyl; 1.				
DR	Pfam; PF00905; Transpeptidase; 1.				
DR	ProDom; PD001895; Transglycosyl; 1.				
DR	SEQUENCE 719 AA; 79758 MW; 5BD397E83B4B3AA6 CRC64;				
QY	1	KIYDNKNQIADLGSERRVNAQNDIPTLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60			
DB	54	KIYDNKNQIADLGSERRVNAQNDIPTLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113			

Query Match 100.0%; Score 3484; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 2.1e-181;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	61	LOSNSLOGGSLTLOOLIKLTVEFSTSPQSTISRAOEAAMLAIOLEOKATOEILFTYINK	120
Db	114	LOSNSLOGGSLTLOOLIKLTVEFSTSPQSTISRAOEAAMLAIOLEOKATOEILFTYINK	173
QY	121	YVMSNGNTGMOTAAQNTYVKDNLNLSLPQALLAGMPQAPNODPYSHPEAADRRNLVYL	180
Db	174	YVMSNGNTGMOTAAQNTYVKDNLNLSLPQALLAGMPQAPNODPYSHPEAADRRNLVYL	233
QY	181	SEMKNOGYSIAEOYEKANPPTITGLOSLSKASNYPRAYMONTLYKEVINQVBEETGYNLLT	240
Db	234	SEMKNOGYSIAEOYEKANPPTITGLOSLSKASNYPRAYMONTLYKEVINQVBEETGYNLLT	293
QY	241	TGMDVYTVNDDEAOKHLMDIYNTDEYVAYAPDDELQVASTIVDYSNGKVIYQLAGARHOSN	300
Db	294	TGMDVYTVNDDEAOKHLMDIYNTDEYVAYAPDDELQVASTIVDYSNGKVIYQLAGARHOSN	353
QY	301	VSEFGINQAVENRBMGSTMKPTITDYAALAEIYGYDSTATIYHDEPRYPGNTNTPYVNMDR	360
Db	354	VSEFGINQAVENRBMGSTMKPTITDYAALAEIYGYDSTATIYHDEPRYPGNTNTPYVNMDR	413
QY	361	GYFGNITLQVYALQOSRNPVAVETLNKGLNRAKFTFLNGLGIDPYSIHYSAISSNTTESD	420
Db	414	GYFGNITLQVYALQOSRNPVAVETLNKGLNRAKFTFLNGLGIDPYSIHYSAISSNTTESD	473
QY	421	KKYGASSEKMAAAYAAPANGSTYKKPMYIHKVYFSDSEKEFEFSNVGTRAKETTYAVMMTD	480
Db	474	KKYGASSEKMAAAYAAPANGSTYKKPMYIHKVYFSDSEKEFEFSNVGTRAKETTYAVMMTD	533
QY	481	MMKTYLVYTCGRNRYLWMLPQAGTGTGNTDEFEIENHITKTSQVAPADELFACTYTRKYSN	540
Db	534	MMKTYLVYTCGRNRYLWMLPQAGTGTGNTDEFEIENHITKTSQVAPADELFACTYTRKYSN	593
QY	541	AVMTGYSNRLPTLVYNGSLTYAAKYVRSMYTLSEGSNPEDMNIPEGLYRNGEEVVFNGAR	600
Db	594	AVMTGYSNRLPTLVYNGSLTYAAKYVRSMYTLSEGSNPEDMNIPEGLYRNGEEVVFNGAR	653
QY	601	STWNSPAPQOPPTRESSSSSSDSTSSQSSSTTSTNNSTTNNNNNTTQOSNTTTPDOONON	660
Db	654	STWNSPAPQOPPTRESSSSSSDSTSSQSSSTTSTNNSTTNNNNNTTQOSNTTTPDOONON	713
QY	661	POPAP 666	
Db	714	POPAP 719	
RESULT	2		
Q9RETR9	ID	PRELIMINARY;	PRT; 719 AA.
Q9RETR9	Q9RETR9	PRELIMINARY;	PRT; 719 AA.
AC	Q9RETR9		
DT	01-MAY-2000 (TREMblrel. 13, Created)		
DT	01-MAY-2000 (TREMblrel. 13, Last sequence update)		
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A.		
GN	PBP1A.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
OC	Streptococcus.		
OC	NCBI_Taxid=1313;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN-SP 1513;		
RC	MEDLINE-21432820; PubMed-11549185;		
RA	Ferroni A., Berche P.;		
RT	"Alterations to penicillin-binding proteins 1A, 2B and 2X amongst		
RT	penicillin-resistant clinical isolates of Streptococcus pneumoniae		
RT	serotype 23F from the nasopharyngeal flora of children."		
RL	J. Med. Microbiol. 50:828-832(2001).		
RL	EMBL: AF210746; AAF17256.1; -		
DR	InterPro: IPR001264; Transglycosyl.		
DR	InterPro: IPR001460; Transpeptidase.		
DR	Pfam: PF00912; Transglycosyl. 1.		
DR	Pfam: PF00905; Transpeptidase; 1.		
DR	ProDom: PD001895; Transglycosyl.1.		

Seq	Sequence	719 AA	79830 MW	5f6776b8bdeae3840 CRC64
Query Match	99.5%	Score 3468	DB 2	Length 719
Best Local Similarity	99.5%	Pred. No. 1.6e-180		
Matches 663	Conservative 2	Mismatches 1	Indels 0	Gaps 0
QY	1	KTYDNKQNLINDLGCERRVNAQANDIPDVLKATYISIEDHREFDRHGDTITILGAFLEKN	60	
DB	54	KTYDNKQNLINDLGCERRVNAQANDIPDVLKATYISIEDHREFDRHGDTITILGAFLEKN	113	
QY	61	LQSNLSGGSTLTQOLIKLTVFSTSTPDQTSRKAKQEAFLAQLQEQKATKQELITVYINK	120	
DB	114	LQSNLSGGSTLTQOLIKLTVFSTSTPDQTSRKAKQEAFLAQLQEQKATKQELITVYINK	173	
QY	121	YVMSNGNMGQTAQNTYGGKDLNNISLPLALLAGMPQAPNOYDPSHPEAAQDRNNLV	180	
DB	174	YVMSNGNMGQTAQNTYGGKDLNNISLPLALLAGMPQAPNOYDPSHPEAAQDRNNLV	233	
QY	181	SEMKNOGYISAEQYKAVNTPTDGLQSLKSAQNPAYMDNLYLKEVINQVEEETGYNLLT	240	
DB	234	SEMKNOGYISAEQYKAVNTPTDGLQSLKSAQNPAYMDNLYLKEVINQVEEETGYNLLT	293	
QY	241	TGMDYTVNVDOEAQKHLMDIYNFDEVAYPPDELOVASTIVDVSNGKYIAQIGARRHSSN	300	
DB	294	TGMDYTVNVDOEAQKHLMDIYNFDEVAYPPDELOVASTIVDVSNGKYIAQIGARRHSSN	353	
QY	301	VSFQINQAVENNRDMGSTMKPTDQAPALEGVYDSTATIVHDEPYNPGTTPVYNNMDR	360	
DB	354	VSFQINQAVENNRDMGSTMKPTDQAPALEGVYDSTATIVHDEPYNPGTTPVYNNMDR	413	
QY	361	GYFGNITLQVLAQOQSRNPVAVETLKKVGLNRAKPTFLNGIDGPSIHYSNAISSNTTESD	420	
DB	414	GYFGNITLQVLAQOQSRNPVAVETLKKVGLNRAKPTFLNGIDGPSIHYSNAISSNTTESD	473	
QY	421	KRYGASSSEKMAAAVAAAFANGSTYYRPMYIHKVYFSDGSEKESNNGTRAMKETTAYMMTD	480	
DB	474	KRYGASSSEKMAAAVAAAFANGSTYYRPMYIHKVYFSDGSEKESNNGTRAMKETTAYMMTD	533	
QY	481	MMKTYLVTGTRNRYLAWLPQAGKGTGNTYDEELENNIKTSQFAPELFGYTRYKSM	540	
DB	534	MMKTYLVTGTRNRYLAWLPQAGKGTGNTYDEELENNIKTSQFAPELFGYTRYKSM	593	
QY	541	AVMTGYSNRLTPLVONGILTVAKYRSMYVLSGGSNPEDMNIPEGLYRNGEFVKNGAR	600	
DB	594	AVMTGYSNRLTPLVONGILTVAKYRSMYVLSGGSNPEDMNIPEGLYRNGEFVKNGAR	653	
QY	601	STWNSPAPQOPPTRESSSSSDSSSTPSQSSSTPTNNSTNNPNNNTTPOOQONTTPOOQON	660	
DB	654	STWNSPAPQOPPTRESSSSSDSSSTPSQSSSTPTNNSTNNPNNNTTPOOQONTTPOOQON	713	
QY	661	POPAQP 666		
DB	714	POPAQP 719		
RESULT	3			
ID	Q9RET3	PRELIMINARY	PRT	719 AA
AC	Q9RET3			
DT	01-MAY-2000 (Tremblrel. 13, Created)			
DE	01-MAY-2000 (Tremblrel. 13, Last sequence update)			
DE	01-DEC-2001 (Tremblrel. 19, Last annotation update)			
OS	PENICILLIN-BINDING PROTEIN IA.			
OS	GN			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus			
OX	NCBI_TaxID=1313;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BM 4200;			
XC	MEDLINE=21432820; Pubmed=11549185;			

RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1a, 2b and 2x amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL; AF210752; AAF17262.1; -.
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00905; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR Prodom; PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79781 MW; DB5993F6BD47F72D CRC64;

Query Match 97.7%; Score 3403; DB 2; Length 719;
Best Local Similarity 97.3%; Pred. No. 5.4e-177;

Matches 648; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIYDKKNOLIADLGSERVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
DB 54 KIYDKKNOLIADLGSERVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAEAMLAIQLEOKATKOEILTYIYNK 120
DB 114 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAEAMLAIQLEOKATKOEILTYIYNK 173
QY 121 YVMSNGNGMGTAAONYYGKDLNNLSLPOLALLAGMPAPNOYDXYSHPEAAODRRNLVL 180
DB 174 YVMSNGNGMGTAAONYYGKDLNNLSLPOLALLAGMPAPNOYDXYSHPEAAODRRNLVL 233
QY 181 SEMKNOGYTSAEOYKAVNTPITDGLQSLKSASNPAYMDNLYKVINOVEEETYNLIT 240
DB 234 SEMKNOGYTSAEOYKAVNTPITDGLQSLKSASNPAYMDNLYKVINOVEEETYNLIT 293
QY 241 TGMDDYTVNDQEAOKHMLDIYNTDEVAYPDDELQVASTIVDVNSGKVIQAOLGARHOSN 300
DB 294 TGMDDYTVNDQEAOKHMLDIYNTDEVAYPDDELQVASTIVDVNSGKVIQAOLGARHOSN 353
QY 301 VSEGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIYHDEBYNTPGINTPYVNMNR 360
DB 354 VSEGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIYHDEBYNTPGINTPYVNMNR 413
QY 361 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKPTFLNGIDYPSIHSNAISSWTESD 420
DB 414 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKPTFLNGIDYPSIHSNAISSWTESD 473
QY 421 KKYGASSEKMAAAYAAAFANGGTYYKPMYIHKVYFSDGSEKESNVTGRAMKETTAYMMTD 480
DB 474 KKYGASSEKMAAAYAAAFANGGTYYKPMYIHKVYFSDGSEKESNVTGRAMKETTAYMMTD 533
QY 481 MKKTYLVLTGTRNAYLAMPQAGKTGTSNYTDEELENHIKTSQFYAPDELFRAGYTRKYSM 540
DB 534 MKKTYLVLTGTRNAYLAMPQAGKTGTSNYTDEELENHIKTSQFYAPDELFRAGYTRKYSM 593
QY 541 AVMTGYSNRLTPLYGNGLTVAAYKYRSMMTYLSBGSNPEDMNIPGLTRNGEFVYKKNAR 600
DB 594 AVMTGYSNRLTPLYGNGLTVAAYKYRSMMTYLSBGSNPEDMNIPGLTRNGEFVYKKNAR 653
QY 601 STWNSPAPQPPSTRESSSSSDSSSTSPSTNNSTTNPNNNTQOOSNTTTPDOQON 660
DB 654 STWNSPAPQPPSTRESSSSSDSSSTSPSTNNSTTNPNNNTQOOSNTTTPDOQON 713
QY 661 POPAP 666
DB 714 POPAP 719

RESULT 4
Q9WVMO PRELIMINARY; PRT; 719 AA.
AC Q9WVMO;
DT 01-NOV-1999 (TREMblrel. 12; Created)
DT 01-NOV-1999 (TREMblrel. 12; Last sequence update)

DT 01-OCT-2001 (TREMblrel. 18; Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN BPPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
RX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PO-342, PO-273, AND PO-341;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
clone of Streptococcus pneumoniae arose by large recombinational
replacements of the cpsA-bppla region.";
RL Microbiology 0:0-0(1999).
DR EMBL; AF139886; AAD43069.1; -.
DR EMBL; AF139884; AAD43067.1; -.
DR EMBL; AF139885; AAD43068.1; -.
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00905; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR Prodom; PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79622 MW; 7EBA6A75EA8FF8B3 CRC64;

Query Match 95.6%; Score 3329; DB 2; Length 719;
Best Local Similarity 94.4%; Pred. No. 5.6e-173;
Matches 629; Conservative 25; Mismatches 12; Indels 0; Gaps 0;

QY 1 KIYDKKNOLIADLGSERVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
DB 54 KIYDKKNOLIADLGSERVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAEAMLAIQLEOKATKOEILTYIYNK 120
DB 114 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAEAMLAIQLEOKATKOEILTYIYNK 173
QY 121 YVMSNGNGMGTAAONYYGKDLNNLSLPOLALLAGMPAPNOYDXYSHPEAAODRRNLVL 180
DB 174 YVMSNGNGMGTAAONYYGKDLNNLSLPOLALLAGMPAPNOYDXYSHPEAAODRRNLVL 233
QY 181 SEMKNOGYTSAEOYKAVNTPITDGLQSLKSASNPAYMDNLYKVINOVEEETYNLIT 240
DB 234 SEMKNOGYTSAEOYKAVNTPITDGLQSLKSASNPAYMDNLYKVINOVEEETYNLIT 293
QY 241 TGMDDYTVNDQEAOKHMLDIYNTDEVAYPDDELQVASTIVDVNSGKVIQAOLGARHOSN 300
DB 294 TGMDDYTVNDQEAOKHMLDIYNTDEVAYPDDELQVASTIVDVNSGKVIQAOLGARHOSN 353
QY 301 VSEGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIYHDEBYNTPGINTPYVNMNR 360
DB 354 VSEGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIYHDEBYNTPGINTPYVNMNR 413
QY 361 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKPTFLNGIDYPSIHSNAISSWTESD 420
DB 414 AFGNITLQYALQOOSRNPAVETLNKVGILNRAKPTFLNGIDYPSIHSNAISSWTESD 473
QY 421 KKYGASSEKMAAAYAAAFANGGTYYKPMYIHKVYFSDGSEKESNVTGRAMKETTAYMMTD 480
DB 474 KKYGASSEKMAAAYAAAFANGGTYYKPMYIHKVYFSDGSEKESNVTGRAMKETTAYMMTD 533
QY 481 MKKTYLVLTGTRNAYLAMPQAGKTGTSNYTDEELENHIKTSQFYAPDELFRAGYTRKYSM 540
DB 534 MKKTYLVLTGTRNAYLAMPQAGKTGTSNYTDEELENHIKTSQFYAPDELFRAGYTRKYSM 593
QY 541 AVMTGYSNRLTPLYGNGLTVAAYKYRSMMTYLSBGSNPEDMNIPGLTRNGEFVYKKNAR 600
DB 594 AVMTGYSNRLTPLYGNGLTVAAYKYRSMMTYLSBGSNPEDMNIPGLTRNGEFVYKKNAR 653
QY 601 STWNSPAPQPPSTRESSSSSDSSSTSPSTNNSTTNPNNNTQOOSNTTTPDOQON 660
DB 654 STWNSPAPQPPSTRESSSSSDSSSTSPSTNNSTTNPNNNTQOOSNTTTPDOQON 713

```
QY 661 POPAP 666
Db 714 POPAP 719

RESULT 5
Q9RET8 PRELIMINARY; PRT; 719 AA.
ID Q9RET8
AC Q9RET8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae.
OC Streptococcus.
OX NCBI_taxid=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SP 1465;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
RT serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL: AF210747; AAF17257.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF009012; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79662 MW; A7B96EA15434A805 CRC64;

Query Match 95.1%; Score 3314; DB 2; Length 719;
Best Local Similarity 94.1%; Pred. No. 3,7e-172;
Matches 627; Conservative 27; Mismatches 12; Indels 0; Gaps 0;
```

```
QY 481 MKKTVLAVGTGRNAYLAWLPQAGKTGTSNYTDEEIEHNKTSQEVAPDELFAGYTRRYSM 540
Db 534 MKKTVLAVGTGRNAYLAWLPQAGKTGTSNYTDEEIEHNKTSQEVAPDELFAGYTRRYSM 593
QY 541 AVWTGYSNRLTPLVGNGITVAAKYRRSMYTLSSGSPNEDMNTEBGLYRNEEFKKGAR 600
Db 594 AVWTGYSNRLTPLVGNGITVAAKYRRSMYTLSSGSPNEDMNTEBGLYRNEEFKKGAR 653
QY 601 STMSPAPQOPSTESSSSSDSSSTPSTNNSTTNPNNNOQSNWTFPDQONON 660
Db 654 STMSPAPQOPSTESSSSSDSSSTPSTNNSTTNPNNNOQSNWTFPDQONON 713
QY 661 POPAP 666
Db 714 POPAP 719
```

```
RESULT 6
Q9WM11 PRELIMINARY; PRT; 719 AA.
ID Q9WM11
AC Q9WM11;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae.
OC Streptococcus.
OX NCBI_taxid=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-URU-E159; M134, URU-E135, AND URU-E157;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 145:0-0(1999).
DR EMBL: AF139889; AAD43073.1; -
DR EMBL: AF139887; AAD43070.1; -
DR EMBL: AF139888; AAD43071.1; -
DR EMBL: AF139889; AAD43072.1; -
DR EMBL: AF159448; AAD48430.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79701 MW; 3EDC8A21C94DAC87 CRC64;
```

```
Query Match 94.7%; Score 3299; DB 2; Length 719;
Best Local Similarity 93.4%; Pred. No. 2,4e-171;
Matches 622; Conservative 30; Mismatches 14; Indels 0; Gaps 0;

QY 1 KITDNKQNLADIGSERVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTRITLGAFLRN 60
Db 54 KITDNKQNLADIGSERVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTRITLGAFLRN 113
QY 61 LQNSLGGSTLTQOLIKLTFSTSDQTSRKAQEAWLAIQLEOKATQOEILTYINK 120
Db 114 LQNSLGGSTLTQOLIKLTFSTSDQTSRKAQEAWLAIQLEOKATQOEILTYINK 173
QY 121 VYNSNGNYGMQTAQNYGKDLNLSLPQALLAGMPOAPNOYDPYSHPEAODRRNLV 180
```

```
DB 174 VYMSNGNYGMQTAOONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHEPAADRNNLV 233
OY 181 SEMKNOGYISAEQYKAVNTPITDGLQSLKASNPAYMDNYLKEVINOVEETGYNLJT 240
DB 234 SEMKNOGYISAEQYKAVNTPITDGLQSLKASNPAYMDNYLKEVINOVEETGYNLJT 293
OY 241 TGMADVNTNDQEAOKHLMIDYNTDEVYAAPDDELQVASTIVVDSNGKVTALGARRQSSN 300
DB 294 TGMADVNTNDQEAOKHLMIDYNTDEVYAAPDDELQVASTIVVDSNGKVTALGARRQSSN 353
OY 301 VSEFGINQAVETNRDMGSKPTTDVAPALEYGYDSTATIVHDEPNYFGTNTPVYNNMR 360
DB 354 VSEFGINQAVETNRDMGSKPTTDVAPALEYGYDSTATIVHDEPNYFGTNTPVYNNMR 413
OY 361 GYFGNTLTQYALQOQSRNVAVETLNKVGJNRKATFLNGIGIDYPSIHYSNATISSNTTESD 420
DB 414 AFGNTLTQYALQOQSRNVAVETLNKVGJNRKATFLNGIGIDYPSIHYSNATISSNTTESD 473
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKYVSDGSEKESNNVGTGRAMKETATAMMD 480
DB 474 KRYGASSEKMAAAYAFANGGTYRKPMYIHKYVSDGSEKESNNVGTGRAMKETATAMMD 533
OY 481 MKKTYLTYGTGRNAYLAMLPOAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
DB 534 MKKTYLTYGTGRGAYLPMLEPOAGKTGTSNYTDEIEIENHIKNTGYVAPDEMFYGYTRKYSM 593
OY 541 AVMTGSNRLTPLYGNGLVAAKYVRSMMTYLSEGSNPEDMNPBELYRNGEYFVFNNGAR 600
DB 594 AVMTGSNRLTPLYGNGLVAAKYVRSMMTYLSEGHPEMDMNPDELFRNGEYFVFNNGAR 653
OY 601 STWNSPAPQOPSTESSSSSDSSTQSSSTPTSTNNSTTTNNNTTQOQNTTTPDOQON 660
DB 654 STWNSPAPQOPSTESSSSSDSSTQSSSTPTSTNNSTTTNNNTTQOQNTTTPDOQON 713
OY 661 POPAOP 666
DB 714 POPAOP 719

RESULT 7
O9RETA PRELIMINARY; PRT; 719 AA.
AC 09RETA;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SP 22861;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1a, 2b and 2x amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL: AF210751; AAF17261.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR ProDom: PD001895; Transglycosyl. 1.
SQ SEQUENCE 719 AA; 79567 MW; 786EDCCBB8BCF286 CRC64;
```

Query Match 94.0%; Score 3274; DB 2; Length 719;
Best Local Similarity 92.9%; Pred. No. 5.5e-170;
Matches 619; Conservative 31; Mismatches 16; Indels 0; Gaps 0;

```
OY 1 KIYDNKNOLIADGSEERRVNAQANDIPTDLVKATVISEDHRRFPDHGIDPTIRILGAFARN 60
DB 54 KIYDNKNOLIADGSEERRVNAQANDIPTDLVKATVISEDHRRFPDHGIDISIRILGAFARN 113
OY 61 LOSNSLQGGSTLTQOLIKLTYESTSTSDQTSRKAQEWALAIQLEKATKQELITVYINK 120
DB 114 LOSNSLQGGSTLTQOLIKLTYESTSTSDQTSRKAQEWALAIQLEKATKQELITVYINK 173
OY 121 VYMSNGNYGMQTAOONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHEPAADRNNLV 180
DB 174 VYMSNGNYGMQTAOONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHEPAADRNNLV 233
OY 181 SEMKNOGYISAEQYKAVNTPITDGLQSLKASNPAYMDNYLKEVINOVEETGYNLJT 240
DB 234 SEMKNOGYISAEQYKAVNTPITDGLQSLKASNPAYMDNYLKEVINOVEETGYNLJT 293
OY 241 TGMADVNTNDQEAOKHLMIDYNTDEVYAAPDDELQVASTIVVDSNGKVTALGARRQSSN 300
DB 294 TGMADVNTNDQEAOKHLMIDYNTDEVYAAPDDELQVASTIVVDSNGKVTALGARRQSSN 353
OY 301 VSEFGINQAVETNRDMGSKPTTDVAPALEYGYDSTATIVHDEPNYFGTNTPVYNNMR 360
DB 354 VSEFGINQAVETNRDMGSKPTTDVAPALEYGYDSTATIVHDEPNYFGTNTPVYNNMR 413
OY 361 GYFGNTLTQYALQOQSRNVAVETLNKVGJNRKATFLNGIGIDYPSIHYSNATISSNTTESD 420
DB 414 AFGNTLTQYALQOQSRNVAVETLNKVGJNRKATFLNGIGIDYPSIHYSNATISSNTTESD 473
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKYVSDGSEKESNNVGTGRAMKETATAMMD 480
DB 474 KRYGASSEKMAAAYAFANGGTYRKPMYIHKYVSDGSEKESNNVGTGRAMKETATAMMD 533
OY 481 MKKTYLTYGTGRNAYLAMLPOAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
DB 534 MKKTYLTYGTGRGAYLPMLEPOAGKTGTSNYTDEIEIENHIKNTGYVAPDEMFYGYTRKYSM 593
OY 541 AVMTGSNRLTPLYGNGLVAAKYVRSMMTYLSEGSNPEDMNPBELYRNGEYFVFNNGAR 600
DB 594 AVMTGSNRLTPLYGNGLVAAKYVRSMMTYLSEGHPEMDMNPDELFRNGEYFVFNNGAR 653
OY 601 STWNSPAPQOPSTESSSSSDSSTQSSSTPTSTNNSTTTNNNTTQOQNTTTPDOQON 660
DB 654 STWNSPAPQOPSTESSSSSDSSTQSSSTPTSTNNSTTTNNNTTQOQNTTTPDOQON 713
OY 661 POPAOP 666
DB 714 POPAOP 719

RESULT 8
O57114 PRELIMINARY; PRT; 719 AA.
AC 057114;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CS111 (CEPHALOSPORIN-RESISTANT CLINICAL ISOLATE);
RX MEDLINE=96012191; PubMed=7574521;
RA Coffey T.J., Daniels M., McDougall L.K., Dowson C.G., Tenover F.C., Spratt B.G.;
RT "Genetic analysis of clinical isolates of Streptococcus pneumoniae with high-level resistance to expanded-spectrum cephalosporins.";
RL Antimicrob. Agents Chemother. 39:1306-1313(1995).
DR EMBL: Z49095; CA88918.1; -
```


DR EMBL; 249094; CAA88917.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR Prodom: PD001895; Transglycosyl; 1.
DR SEQUENCE 719 AA; 79712 MW; 5396607C0E67D06E CRC64;

Query Match 93.2%; Score 3247; DB 2; Length 719;
Best Local Similarity 91.9%; Pred. No. 1.6e-166;
Matches 612; Conservative 33; Mismatches 21; Indels 0; Gaps 0;

```
QY 1 KIYDNKNLIADLIGSERVNAQANDIPTDLVKAIVSIEDHREFDRGIDTIRILGAFLRN 60
DB 54 KIYDNKNLIADLIGSERVNAQANDIPTDLVKAIVSIEDHREFDRGIDTIRILGAFLRN 113
QY 61 LOSNLSGGSTLTQOLIKLTYFSTSDQTSRKAQEMALAIQEQKATKQELITYYINK 120
DB 114 LOSNLSGGSTLTQOLIKLTYFSTSDQTSRKAQEMALAIQEQKATKQELITYYINK 173
QY 121 VYMSNGNGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDYSHPPEAQRRLVYL 180
DB 174 VYMSNGNGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDYSHPPEAQRRLVYL 233
QY 181 SEMKNGYISAQYKAVNPTITDGLQSLKSASNPAYMDNLYKQVINOVEETGYNLLT 240
DB 234 SEMKNGYISAQYKAVNPTITDGLQSLKSASNPAYMDNLYKQVINOVEETGYNLLT 293
QY 241 TGMDEVYTVNDOEAKHLMIDYNTDEVYVAPDELOVASTYDVNSGKYIAQLGARHQSNN 300
DB 294 TGMDEVYTVNDOEAKHLMIDYNSDQYVSPDDDLQVASTYDVNSGKYIAQLGARHQSNN 353
QY 301 VSFGINQAVETNRDGMGSMKPTTDYAPALEGVYSTATVYDEPYNPGTNTPYNNMNR 360
DB 334 VSFGINQAVETNRDGMGSMKPTTDYAPALEGVYSTATVYDEPYNPGTNTPLYNMNDH 413
QY 361 GYFGNITLQYALQOQRNRYPAVETLNKVGILNRAKTFPLNGIGIDYPSIHSNMAISSWTESD 420
DB 414 VYFGNITLQYALQOQRNRYPAVETLNKVGILNRAKTFPLNGIGIDYPSIHSNMAISSWTESN 473
QY 421 KKYGASSEKMAAAYAAPANGGTYKPKYIHKVYSDGSEKESNVGTRAMKETTAYMMTD 480
DB 474 KKYGASSEKMAAAYAAPANGGTYKPKYIHKVYSDGSEKESNVGTRAMKETTAYMMTE 533
QY 481 MKKTVLTGGRNAYLAFLPAQAGKTGTSNYTDEETENHKTQFAPARPELAGYRRTKSM 540
DB 534 MKKTVLTGGRNAYLAFLPAQAGKTGTSNYTDEETENHKTQFAPARPELAGYRRTKSM 593
QY 541 AVMTGYSNRLTPLYNGGLTVAAKYRSMNTYLSGGSNPEDNNIPGLYRNGEFYKNGAR 600
DB 594 AVMTGYSNRLTPLYNGGLTVAAKYRSMNTYLSGGSNPEDNNIPGLYRNGEFYKNGAR 653
QY 601 STWNSPAPQPPSTESSSSSDSSTSGSSSTPTSTNNSTTNPNNNTQOQNTTPOQONON 660
DB 654 STWNSPAPQPPSTESSSSSDSSTSGSSSTPTSTNNSTTNPNNNTQOQNTTPOQONON 713
QY 661 POPAP 666
DB 714 POPAP 719
```

RESULT 9
Q9RET6 PRELIMINARY; PRT; 719 AA.
AC Q9RET6;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PB1A.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP 1053;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
RT serotype 23F from the nasopharyngeal flora of children."
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL; AF210749; AAF1259.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR Prodom: PD001895; Transglycosyl; 1.
DR SEQUENCE 719 AA; 79638 MW; 1DAE3C937048E16 CRC64;

Query Match 93.2%; Score 3246; DB 2; Length 719;
Best Local Similarity 92.0%; Pred. No. 1.8e-166;
Matches 613; Conservative 35; Mismatches 18; Indels 0; Gaps 0;

```
QY 1 KIYDNKNLIADLIGSERVNAQANDIPTDLVKAIVSIEDHREFDRGIDTIRILGAFLRN 60
DB 54 KIYDNKNLIADLIGSERVNAQANDIPTDLVKAIVSIEDHREFDRGIDTIRILGAFLRN 113
QY 61 LOSNLSGGSTLTQOLIKLTYFSTSDQTSRKAQEMALAIQEQKATKQELITYYINK 120
DB 114 LOSNLSGGSTLTQOLIKLTYFSTSDQTSRKAQEMALAIQEQKATKQELITYYINK 173
QY 121 VYMSNGNGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDYSHPPEAQRRLVYL 180
DB 174 VYMSNGNGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDYSHPPEAQRRLVYL 233
QY 181 SEMKNGYISAQYKAVNPTITDGLQSLKSASNPAYMDNLYKQVINOVEETGYNLLT 240
DB 234 SEMKNGYISAQYKAVNPTITDGLQSLKSASNPAYMDNLYKQVINOVEETGYNLLT 293
QY 241 TGMDEVYTVNDOEAKHLMIDYNTDEVYVAPDELOVASTYDVNSGKYIAQLGARHQSNN 300
DB 294 TGMDEVYTVNDOEAKHLMIDYNSDQYVSPDDDLQVASTYDVNSGKYIAQLGARHQSNN 353
QY 301 VSFGINQAVETNRDGMGSMKPTTDYAPALEGVYSTATVYDEPYNPGTNTPYNNMNR 360
DB 334 VSFGINQAVETNRDGMGSMKPTTDYAPALEGVYSTATVYDEPYNPGTNTPLYNMNDH 413
QY 361 GYFGNITLQYALQOQRNRYPAVETLNKVGILNRAKTFPLNGIGIDYPSIHSNMAISSWTESD 420
DB 414 AVFGNITLQYALQOQRNRYPAVETLNKVGILNRAKTFPLNGIGIDYPSIHSNMAISSWTESN 473
QY 421 KKYGASSEKMAAAYAAPANGGTYKPKYIHKVYSDGSEKESNVGTRAMKETTAYMMTD 480
DB 474 KKYGASSEKMAAAYAAPANGGTYKPKYIHKVYSDGSEKESNVGTRAMKETTAYMMTE 533
QY 481 MKKTVLTGGRNAYLAFLPAQAGKTGTSNYTDEETENHKTQFAPARPELAGYRRTKSM 540
DB 534 MKKTVLTGGRNAYLAFLPAQAGKTGTSNYTDEETENHKTQFAPARPELAGYRRTKSM 593
QY 541 AVMTGYSNRLTPLYNGGLTVAAKYRSMNTYLSGGSNPEDNNIPGLYRNGEFYKNGAR 600
DB 594 AVMTGYSNRLTPLYNGGLTVAAKYRSMNTYLSGGSNPEDNNIPGLYRNGEFYKNGAR 653
QY 601 STWNSPAPQPPSTESSSSSDSSTSGSSSTPTSTNNSTTNPNNNTQOQNTTPOQONON 660
DB 654 STWNSPAPQPPSTESSSSSDSSTSGSSSTPTSTNNSTTNPNNNTQOQNTTPOQONON 713
QY 661 POPAP 666
DB 714 POPAP 719
```


RESULT 10
054947 PRELIMINARY: PRT: 719 AA.
AC 054949:
DB 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56742;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT *Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain";
RL EMBL J. 11:3831-3836(1992).
DR EMBL; X67869; CAA48069.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA: 79750 MW: 489B6C6E0CB861CD CRC64;

Query Match 93.2%; Score 3246; DB 2; Length 719;
Best Local Similarity 92.0%; Pred. No. 1,8e-168;
Matches 613; Conservative 31; Mismatches 22; Indels 0; Gaps 0;
DB 1 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKATVSTEDHREFDHRGIDPRTILGAFNRN 60
DB 54 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKATVSTEDHREFDHRGIDSRILGAFNRN 113
DB 61 LOSNSLQSGSTLTQOLIKLTYFSTSDQTSRKAEAWMLAIOLEKATKQELITYYINK 120
DB 114 LOSNSLQSGSTLTQOLIKLTYFSTSDQTSRKAEAWMLAIOLEKATKQELITYYINK 173
DB 121 VYMSNGNYGMQTAQNYVYKDLNNLSLPQALLAGMPQAPNOYDPYSHPEAADRNRNLY 180
DB 174 VYMSNGNYGMQTAQNYVYKDLNNLSLPQALLAGMPQAPNOYDPYSHPEAADRNRNLY 233
DB 181 SEKNNGYISAEQYERKAVNPTITDGLQSLKSASNYPAYMDNYLKEVYINOVEEETGYNLT 240
DB 234 SEKNNGYISAEQYERKAVNPTITDGLQSLKSASNYPAYMDNYLKEVYINOVEEETGYNLT 293
DB 241 TGMDEVYTNVDOEAKHLMDIYNTDEVYAYPDDDELQVASTIVDSNGKVIQAOLGARHQS 300
DB 294 TGMDEVYTNVDOEAKHLMDIYNTDEVYAYPDDDELQVASTIVDSNGKVIQAOLGARHQS 353
DB 301 VSFNGINQAVETNRDQSGTAKPTIDYAPALEYGYDSTATIVHDEPNYPCGTNTPYVNM 360
DB 354 VSFNGINQAVETNRDQSGTAKPTIDYAPALEYGYDSTATIVHDEPNYPCGTNTPYVNM 413
DB 361 GYFGNTTLQYALQOQSRNVPAVETLNKVGINRAKTFPLNGIGIDVPSIHYSAISNTTES 420
DB 414 ATFGNTTLQYALQOQSRNVPAVETLNKVGINRAKTFPLNGIGIDVPSIHYSAISNTTES 473
DB 421 KYVGASSEKMAAAYAFANGGTYYKPMYIHKVYVSDGSEKESFNGVTRAMKETATAYMMD 480
DB 474 KYVGASSEKMAAAYAFANGGTYYKPMYIHKVYVSDGSEKESFNGVTRAMKETATAYMMD 533
DB 481 MKATVLTGTGRNAVLAFLPQAGKTGTSYNTDEEINHIKTSQFAPADELFACTRYKYS 540
DB 534 MKATVLTGTGRNAVLAFLPQAGKTGTSYNTDEEINHIKTSQFAPADELFACTRYKYS 593
DB 541 AWTGYSNRLTPLVGOFLVAAKVYRSMYTYSEGSNPEDWNPBGLYRNGEVPFNGAR 600

DB 594 AWTGYSNRLTPLVGOFLVAAKVYRSMYTYSEGSNPEDWNPBGLYRNGEVPFNGAR 653
DB 601 STYNSPAPQOPSTESSSSSDSTSSSTPTSTNNSTTTPNNNTQOQNTTPDQONON 660
DB 654 PINTESTQOQSTAESSSSSSDSTSSSTPTSTNNSTTTPNNNTQOQNTTPDQONON 713
DB 661 POPAOP 666
DB 714 POPAOP 719

RESULT 11
054947 PRELIMINARY: PRT: 719 AA.
AC 054947:
DB 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=681;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT *Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain";
RL EMBL J. 11:3831-3836(1992).
DR EMBL; X67869; CAA48069.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA: 79710 MW: D90EB42D5380D9E1 CRC64;

Query Match 93.1%; Score 3245; DB 2; Length 719;
Best Local Similarity 92.0%; Pred. No. 2,1e-168;
Matches 613; Conservative 35; Mismatches 18; Indels 0; Gaps 0;
DB 1 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKATVSTEDHREFDHRGIDPRTILGAFNRN 60
DB 54 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKATVSTEDHREFDHRGIDSRILGAFNRN 113
DB 61 LOSNSLQSGSTLTQOLIKLTYFSTSDQTSRKAEAWMLAIOLEKATKQELITYYINK 120
DB 114 LOSNSLQSGSTLTQOLIKLTYFSTSDQTSRKAEAWMLAIOLEKATKQELITYYINK 173
DB 121 VYMSNGNYGMQTAQNYVYKDLNNLSLPQALLAGMPQAPNOYDPYSHPEAADRNRNLY 180
DB 174 VYMSNGNYGMQTAQNYVYKDLNNLSLPQALLAGMPQAPNOYDPYSHPEAADRNRNLY 233
DB 181 SEKNNGYISAEQYERKAVNPTITDGLQSLKSASNYPAYMDNYLKEVYINOVEEETGYNLT 240
DB 234 SEKNNGYISAEQYERKAVNPTITDGLQSLKSASNYPAYMDNYLKEVYINOVEEETGYNLT 293
DB 241 TGMDEVYTNVDOEAKHLMDIYNTDEVYAYPDDDELQVASTIVDSNGKVIQAOLGARHQS 300
DB 294 TGMDEVYTNVDOEAKHLMDIYNTDEVYAYPDDDELQVASTIVDSNGKVIQAOLGARHQS 353
DB 301 VSFNGINQAVETNRDQSGTAKPTIDYAPALEYGYDSTATIVHDEPNYPCGTNTPYVNM 360
DB 354 VSFNGINQAVETNRDQSGTAKPTIDYAPALEYGYDSTATIVHDEPNYPCGTNTPYVNM 413
DB 361 GYFGNTTLQYALQOQSRNVPAVETLNKVGINRAKTFPLNGIGIDVPSIHYSAISNTTES 420

```
|||||
Db 414 AYFGMITTQYALQOOSRNTAVETLNKVGIDRAKTEFLNGIDIDYPSMHVANAISNTTESN 473
OY 421 KKYGASSEKMAAAYAFANGGTYKPMYTHKVFSDGSEKESNNGTRAMKETTAYMMTD 480
Db 474 KQYGASSEKMAAAYAFANGGTYKPMYTHKVFSDGSEKESNNGTRAMKETTAYMMTE 533
OY 481 MKKTVLTGTGNAYLAMLPOAGKTGTSNYTDEEIEHNHKTISOFAVAPDELAFAGYTRKYSM 540
Db 534 MKKTVLTGTGNAYLAMLPOAGKTGTSNYTDEEIEHNHKTISOFAVAPDELAFAGYTRKYSM 593
OY 541 AVWTGYSNRLPPLVGNGLTVAAKYRSMYTYLSEGSNPEDMNIPEGLYRNGEVEKNGAR 600
Db 594 AVWTGYSNRLPPLVGNGLTVAAKYRSMYTYLSEGSNPEDMNIPEGLYRNGEVEKNGAR 653
OY 601 STWNSPAPQOQPSTSSSSSDSSTSSSTPSTNNSTTTPNNNTQOOSNTTPDOONON 660
Db 654 STWNSPAPQOQPSTSSSSSDSSTSSSTPSTNNSTTTPNNNTQOOSNTTPDOONON 713
OY 661 POPAP 666
Db 714 POPAP 719

RESULT 12
O54946
ID 054946 PRELIMINARY; PRT: 719 AA.
AC 054946:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA OR PAP1A.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1313;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-456:
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain."
RL EMBO J. 11:3831-3836(1992).
RN 12
RP SEQUENCE OF 310-619 FROM N.A.
RC STRAIN-8303, AND 35193;
RA Smith A.M., Klugman K.P.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
RN 13
RP SEQUENCE OF 320-717 FROM N.A.
RC STRAIN-#17/246, #20/898, AND #27/SHA3;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a thr-371 substitution in a conserved amino acid motif
RT of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae."
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
RN 14
RP SEQUENCE FROM N.A.
RC STRAIN-SP-665;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-Dppla region."
RL Microbiology 0:0-0(1999).
DR EMBL: X67868; CAA48068.1; -
DR EMBL: AF046230; AAC24695.1; -
DR EMBL: AB006876; BAA32072.1; -
DR EMBL: AF139863; AAD43066.1; -
DR InterPro: IPR001264; Transglycosyl.
```

```
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA: 79736 MW: F090F91522C834DC CRC64:

Query Match 93.1%; Score 3243; DB 2; Length 719;
Best Local Similarity 91.9%; Pred. No. 2,7e-168;
Matches 612; Conservative 32; Mismatches 22; Indels 0; Gaps 0;

OY 1 KIYDNKNOIADLGSERRVNAQANDIPDLYKAIVSIEDHREDFRGIDITRILGAFLEN 60
Db 54 KIYDNKNOIADLGSERRVNAQANDIPDLYKAIVSIEDHREDFRGIDISIRILGAFLEN 113
OY 61 LOSNLSGGSTLTQOLIKITFYSTSTSDQTSIRKQOEWALQLOKATKQILTYINK 120
Db 114 LOSNLSGGSTLTQOLIKITFYSTSTSDQTSIRKQOEWALQLOKATKQILTYINK 173
OY 121 VYMSNGNYGMOTAAQNYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPEAODRRNLVL 180
Db 174 VYMSNGNYGMOTAAQNYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPEAODRRNLVL 233
OY 181 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNYPAYMDNYLKEVINQVEBETGYNLT 240
Db 234 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNYPAYMDNYLKEVINQVEBETGYNLT 293
OY 241 TGMDDYTVNDQAKHMDIYNTDEVAVAPDELOVASTIVYVNSGKYTAQUGAHHQSSN 300
Db 294 TGMDDYTVNDQAKHMDIYNTDEVAVAPDELOVASTIVYVNSGKYTAQUGAHHQSSN 353
OY 301 VSEGTNOAVETNRDWSITMKPTIDYAPALEYGVYSTATIVHDEPNYPGTNTPYNNMR 360
Db 354 VSEGTNOAVETNRDWSITMKPTIDYAPALEYGVYSTATIVHDEPNYPGTNTPYNNMR 413
OY 361 GYFGNITLQYALQOOSRNTAVETLNKVGIDRAKTEFLNGIDIDYPSIHYSNAISNTTESD 420
Db 414 AYFGMITTQYALQOOSRNTAVETLNKVGIDRAKTEFLNGIDIDYPSMHVANAISNTTESN 473
OY 421 KKYGASSEKMAAAYAFANGGTYKPMYTHKVFSDGSEKESNNGTRAMKETTAYMMTD 480
Db 474 KQYGASSEKMAAAYAFANGGTYKPMYTHKVFSDGSEKESNNGTRAMKETTAYMMTE 533
OY 481 MKKTVLTGTGNAYLAMLPOAGKTGTSNYTDEEIEHNHKTISOFAVAPDELAFAGYTRKYSM 540
Db 534 MKKTVLTGTGNAYLAMLPOAGKTGTSNYTDEEIEHNHKTISOFAVAPDELAFAGYTRKYSM 593
OY 541 AVWTGYSNRLPPLVGNGLTVAAKYRSMYTYLSEGSNPEDMNIPEGLYRNGEVEKNGAR 600
Db 594 AVWTGYSNRLPPLVGNGLTVAAKYRSMYTYLSEGSNPEDMNIPEGLYRNGEVEKNGAR 653
OY 601 STWNSPAPQOQPSTSSSSSDSSTSSSTPSTNNSTTTPNNNTQOOSNTTPDOONON 660
Db 654 STWNSPAPQOQPSTSSSSSDSSTSSSTPSTNNSTTTPNNNTQOOSNTTPDOONON 713
OY 661 POPAP 666
Db 714 POPAP 719

RESULT 13
O54948
ID 054948 PRELIMINARY; PRT: 719 AA.
AC 054948:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
```

NCBI_TaxID=1313;
[1]
SEQUENCE FROM N.A.
RC STRAIN=670;
RX MEDLINE=93010977; Pubmed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMBL J. 11:3831-3836(1992).
DR EMBL: X67867; CAA48067.1;
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR Prodom: PD001895; Transglycosyl. 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA; 79611 MW; 1DBFDA93704930B CRC64;

Query Match 93.0%; Score 3241; DB 2; Length 719;
Best Local Similarity 91.4%; Pred. No. 3,4e-168;
Matches 612; Conservative 36; Mismatches 18; Indels 0; Gaps 0;

1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHFRFDHGRGIDTIRILGAFLRN 60
54 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHFRFDHGRGIDIRILGAFLRN 113
61 LOSNSLGGSGTTLQOLIKLTYFSTSDQTSRKAQEWMLAIOLEKATKQELITYYINK 120
114 LOSNSLGGSGTTLQOLIKLTYFSTSDQTSRKAQEWMLAIOLEKATKQELITYYINK 173
121 VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPPAADRRNLV 180
174 VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPPAADRRNLV 233
181 SEMKNGYISAEQYERKAVNPTITDGLQSLKASNPAYMDNYLKEYINQVEETGYNLT 240
234 SEMKNGYISAEQYERKAVNPTITDGLQSLKASNPAYMDNYLKEYINQVEETGYNLT 293
241 TGMDEVYTNDOEAQKHLMDIYNTDEVYAYPDDELQVASTIVDSNGKVIQAOLGARHOSN 300
294 TGMDEVYTNDOEAQKHLMDIYNSDYVSPDDLOVASTIVDSNGKVIQAOLGARHOSN 353
301 VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYDSTATVHDEPNYVPGTNPVYNMNR 360
354 VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYDSTATVHDEPNYVPGTNPVYNMNR 413
361 GYFGNTITLOALQOOSRNVPAVETLNKGLNRKATFLNGLIDYPSIHYSAISSNTTESD 420
414 AYFGNTITLOALQOOSRNVPAVETLNKGLNRKATFLNGLIDYPSIHYSAISSNTTESN 473
421 KRYGASSEKMAAAYAFANGGYTYPYIHKVYVSDGSEKEFSNVGTRAKETAYAMMD 480
474 KRYGASSEKMAAAYAFANGGYTYPYIHKVYVSDGSEKEFSNVGTRAKETAYAMMD 533
481 MKTIVLYTGTGRNAYLAWLPQAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 540
534 MKTIVLYTGTGRNAYLAWLPQAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 593
541 AVWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFNKNGAR 600
594 AVWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFNKNGAR 653
601 STWNSPAPQOPSTESSSSSDSTSSSTSTSTNNSTTNNNNNTTQOOSNTTPOOONON 660
654 STWNSPAPQOPSTESSSSSDSTSSSTSTSTNNSTTNNNNNTTQOOSNTTPOOONON 713
661 POPAOP 666
714 POPAOP 719

RESULT 14
ID 09RET5 PRELIMINARY; PRT: 719 AA.
AC 09RET5;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DT PENICILLIN-BINDING PROTEIN 1A.
GN ppa1a.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP 1470;
RX MEDLINE=21432820; Pubmed=11549185;
RA Ferion A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
RT serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL: AF210750; AAF17260.1;
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR Prodom: PD001895; Transglycosyl. 1.
SQ SEQUENCE 719 AA; 79734 MW; 3BE77571DBC86496 CRC64;

Query Match 92.7%; Score 3231; DB 2; Length 719;
Best Local Similarity 91.4%; Pred. No. 1.2e-167;
Matches 609; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHFRFDHGRGIDTIRILGAFLRN 60
54 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHFRFDHGRGIDIRILGAFLRN 113
61 LOSNSLGGSGTTLQOLIKLTYFSTSDQTSRKAQEWMLAIOLEKATKQELITYYINK 120
114 LOSNSLGGSGTTLQOLIKLTYFSTSDQTSRKAQEWMLAIOLEKATKQELITYYINK 173
121 VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPPAADRRNLV 180
174 VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPPAADRRNLV 233
181 SEMKNGYISAEQYERKAVNPTITDGLQSLKASNPAYMDNYLKEYINQVEETGYNLT 240
234 SEMKNGYISAEQYERKAVNPTITDGLQSLKASNPAYMDNYLKEYINQVEETGYNLT 293
241 TGMDEVYTNDOEAQKHLMDIYNTDEVYAYPDDELQVASTIVDSNGKVIQAOLGARHOSN 300
294 TGMDEVYTNDOEAQKHLMDIYNSDYVSPDDLOVASTIVDSNGKVIQAOLGARHOSN 353
301 VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYDSTATVHDEPNYVPGTNPVYNMNR 360
354 VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYDSTATVHDEPNYVPGTNPVYNMNR 413
361 GYFGNTITLOALQOOSRNVPAVETLNKGLNRKATFLNGLIDYPSIHYSAISSNTTESD 420
414 AYFGNTITLOALQOOSRNVPAVETLNKGLNRKATFLNGLIDYPSIHYSAISSNTTESN 473
421 KRYGASSEKMAAAYAFANGGYTYPYIHKVYVSDGSEKEFSNVGTRAKETAYAMMD 480
474 KRYGASSEKMAAAYAFANGGYTYPYIHKVYVSDGSEKEFSNVGTRAKETAYAMMD 533
481 MKTIVLYTGTGRNAYLAWLPQAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 540
534 MKTIVLYTGTGRNAYLAWLPQAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 593
541 AVWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFNKNGAR 600

Db	594	AWTGYTSNNLPTVGGDFLVAAKVYRSMITLYLSEDTNHPEDMTMPDGLFRNGEVPFVNGAR	653
Oy	601	STWNSPAPQPPSTSSSSSSSDSSSTPSQSSSTTPSTNNSTTTNPNNNTQQSNTTPDQONON	660
Db	654	PWTPESTTQSSSTAASSSSSSSDSSSTPSQSSSTTPSTNNSTTTNPNNNTQQSNTTPDQONON	713
Oy	661	POPAP 666	
Db	714	POPAP 719	
RESULT	15		
ID	054950		
AC	054950	PRELIMINARY;	PRT: 719 AA.
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, last annotation update)	
DE	PENICILLIN-BINDING PROTEIN_1A (FRAGMENT).		
GN	PONA.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1313;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-2039;		
RX	MEDLINE-93010977; Pubmed-1396576;		
RA	Martin C., Sibold C., Hakenbeck R.;		
RT	"Relatedness of penicillin-binding protein 1a genes from different		
RT	clones of penicillin-resistant Streptococcus pneumoniae isolated in		
RT	South Africa and Spain.";		
RL	EMBO J. 11:3831-3836(1992).		
DR	EMBL, X67670; CAA48070.1;		
DR	InterPro: IPR001264; Transglycosyl.		
DR	InterPro: IPR001460; Transpeptidase.		
DR	Pfam: PF00912; Transglycosyl; 1.		
DR	Pfam: PF00905; Transpeptidase; 1.		
DR	ProDom: PD001895; Transglycosyl; 1.		
FT	NON_TER		
FT	719		
SO	SEQUENCE	719 AA: 79874 MW: FACBFBAB303D865D4 CRC64:	

Query Match	91.6%	Score 3192	DB 2	Length 719
Best Local Similarity	90.4%	Pred. No. 1	6e-165	
Matches 602	Conservative	41	Mismatches 23	Indels 0
Gaps				0
QY	1	KIYDKNQLIADLGSERRNAQANDIPDTLVAAIYSIEDHREFDRHGIDITIRILGAFLRN	60	
Db	54	KIFDKNELIADLGSERRNAQANLPDVLVAIYSIEDHREFDRHCVTIRILGAFLRN	113	
QY	61	LOSNSLGGSTLTQOLIKLTFYFSTSDQTSIRKQAEAWLAIQLEOKATKOEILYYINK	120	
Db	114	LQNNLSLGGSTLTQOLIKLTFYFSTSDQTSIRKQAEAWLAIQLEOKATKOEILYYINK	173	
QY	121	YVMSGNGVMQTAQAAONYGKDLNNISLPOLALLAGHPQAPNOYDPSHPEDAADRNLVL	180	
Db	174	YVMSGNGVMQTAQAAONYGKDLRELISLPOLALLAGHPQAPNOYDPSHPEDAADRNLVL	233	
QY	181	SEMKQOGYISAOYKKAANTPTDGLQSLKSAKNYPAYMDNTLKEVINOVEETGYNLLT	240	
Db	234	SEMKQOGYISAOYKKAANTPTDGLQSLKSAKNYPAYMDNTLKEVIDVEETGYNLLT	293	
QY	241	TGMDVYTVNVDQEAQKHLMDIYNTDEVAAPDEDELQVASTIVVNSNGKYTAQGLGARHSSN	300	
Db	294	TGMDVYTVNVDQEAQKRLMDIYNSDQYVSRPDDDLQVASTIVVNSNGKYTAQGLGARHQA3N	353	
QY	301	VSEGINQAVETNRDQGSTWKPLTDYAPALEYGVYDSTATIVHDEPYNPGTNPVYNMWR	360	
Db	354	VSEGINQAVETNRDQGSTWKPLTDYAPALEYGVYDSTATIVHDPVNPGBTDLPTNMDC	413	
QY	361	GYFGNITLQYALQGRNRPVETLKNKVGKLNRAKKTPLNGIGIDYPSIHSNAISSNTSDE	420	

Db	414	VYFGNITITQVALQOSRNBTAVENTLINTKVLGRKFTLNGLGDVPSMAYANASINNTESN	473
Qy	421	KRYGASSSEKMAAATAAFRANGSTYYKKRPYIKHKVYPSDSESEKFEFNVGTRAKKETTAYMTD	480
Db	474	KOYGASSSEKMAAATAAFRANGSTYYKKRPYIKHKVYPSDSESEKFEFNVGTRAKKETTAYMTD	533
Qy	481	MKRTVLITGTGRNMYLAMLPOAGKTGTSNVTDEIEIENHIKTSOFVADLEFAGYTRKYSM	540
Db	534	MKRTVLITGTGRNMYLAMLPOAGKTGTSNVTDEIEIENHIKTSOFVADLEFAGYTRKYSM	593
Qy	541	AVMVGYSNRLPLVINGLITVAAKYRBSMNYLISGSPEDOMNIPGELYRNGEYFVKNGAR	600
Db	594	AVMVGYSNRLPLVINGLITVAAKYRBSMNYLISGSPEDOMNIPGELYRNGEYFVKNGAR	653
Qy	601	STWNSPAPQCPSTESSSSSSSDSSSTPSTPSTNNSTTTNNPNNNTQOQNTTPDOONON	660
Db	654	STWNSPAPQCPSTESSSSSSSDSSSTPSTPSTNNSTTTDPNNNTQOQNTTPDOONON	713
Qy	661	POPAPQ 666	
Db	714	POPAPQ 719	

Search completed: August 11, 2002, 04:34:19
Job time: 600 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 19:11:54 ; Search time 3654.27 Seconds
(without alignments)
11447.470 Million cell updates/sec

Title: US-08-961-083-1
1999

Perfect score: 1 TAAATCTGACGACATATAA.....ATCCTCAGACGACACACA 1999

Sequence: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Scoring table: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
------------	-------	-------------	--------	----	----	-------------

1	1999	100.0	1999	6	ARI20238	ARI20238 Sequence
2	1999	100.0	10711	6	BD003825	BD003825 Polynucle
3	1999	100.0	11334	1	AE007349	AE007349 Streptoco
4	1992.6	99.7	2160	1	AF210745	AF210745 Streptoco
5	1986.2	99.4	2157	1	SPON8A	SPON8A Streptoco
6	1983	99.2	2160	1	AF210746	AF210746 Streptoco
7	1981.4	99.1	2157	1	SPON7A	SPON7A Streptoco
8	1981.4	99.1	2160	6	AX111315	AX111315 Streptoco
9	1981.4	99.1	3378	1	STRPONA	STRPONA Streptoco
10	1981.4	99.1	10313	1	AE008414	AE008414 Streptoco
11	1887	94.4	2160	1	AF210752	AF210752 Streptoco
12	1704.6	85.3	2160	1	AF139884	AF139884 Streptoco
13	1704.6	85.3	2160	1	AF139885	AF139885 Streptoco
14	1704.6	85.3	2160	1	AF139886	AF139886 Streptoco
15	1677.4	83.9	2160	1	AF139887	AF139887 Streptoco
16	1677.4	83.9	2160	1	AF139888	AF139888 Streptoco
17	1677.4	83.9	2160	1	AF139889	AF139889 Streptoco
18	1677.4	83.9	2160	1	AF139890	AF139890 Streptoco
19	1677.4	83.9	2160	1	AF159448	AF159448 Streptoco
20	1677.4	83.9	2160	6	AX111316	AX111316 Streptoco
21	1645.4	82.3	2157	1	SPONNA4	SPONNA4 Streptoco
22	1643.8	82.2	2157	1	SPONNA3	SPONNA3 Streptoco
23	1643.8	82.2	2160	1	AF210747	AF210747 Streptoco
24	1643.8	82.2	30457	2	SPNEU1914	SPNEU1914 Streptoco
25	1642.2	82.2	2160	1	AF139883	AF139883 Streptoco
26	1642.2	82.2	2160	6	AX110436	AX110436 Streptoco
27	1640.6	82.0	2160	1	AF210750	AF210750 Streptoco
28	1639	82.0	2160	1	AF210751	AF210751 Streptoco
29	1621.4	81.1	2172	1	SPCS1091A	SPCS1091A Streptoco
30	1621.4	81.1	2172	1	SPCS1111A	SPCS1111A Streptoco
31	1621.4	81.1	2172	1	SPCS1111A	SPCS1111A Streptoco
32	1608.6	80.5	2160	1	AF210749	AF210749 Streptoco
33	1607	80.4	2157	1	SPONNA1	SPONNA1 Streptoco
34	1607	80.4	2157	1	SPONNA2	SPONNA2 Streptoco
35	1607	80.4	2157	6	AX111407	AX111407 Streptoco
36	1563.8	78.2	2154	1	AF210748	AF210748 Streptoco
37	1556.2	77.8	1826	1	SPONNA6	SPONNA6 Streptoco
38	1442.2	72.1	2157	1	SPONNA5	SPONNA5 Streptoco
39	1247.4	62.4	1249	6	AX111327	AX111327 Streptoco
40	1185.4	59.3	1195	1	AB006874	AB006874 Streptoco
41	1182.2	59.1	1195	1	AB006868	AB006868 Streptoco
42	1161.4	58.1	2184	1	SMI295852	SMI295852 Streptoco
43	1122.2	56.1	1199	6	AX110281	AX110281 Streptoco
44	1107.8	55.4	2528	1	STRPONA	STRPONA Streptoco
45	1105.4	55.3	1195	1	AB006879	AB006879 Streptoco

ALIGNMENTS

RESULT 1
ARI20238
LOCUS ARI20238 1999 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6159469.
ACCESSION ARI20238
VERSION ARI20238.1 GI:14103814
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1999)

AUTHORS Choi, G.H., Kunsch, C.A., Barash, S.C., Dillon, P.J., Dougherty, B., Fannon, M.R. and Rosen, C.A.
TITLE Streptococcus pneumoniae antigens and vaccines
JOURNAL Patent: US 6159469-A 12-DEC-2000;
FEATURES location/Qualifiers
source 1. 1999

BASE COUNT 656 a 485 c 376 g 482 t
ORIGIN /organism="unknown"

Query Match 100.0%; Score 1999; DB 6; Length 1999;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TAAATCTAGACAAATAAATCAACATCATGTGACTGGGCTCTGAACCCCGCTCAA 60
Db 1 TAAATCTAGACAAATAAATCAACATCATGTGACTGGGCTCTGAACCCCGCTCAA 60
QY 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATGGAAGCA 120
Db 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATGGAAGCA 120
QY 121 TCGCTTCTTGACACACAGGGGGATGATACATCCGTAATCCTGGAGCTTTCTGGCGAA 180
Db 121 TCGCTTCTTGACACACAGGGGGATGATACATCCGTAATCCTGGAGCTTTCTGGCGAA 180
QY 181 TCTGCAAGCAATTCCTCCCAAGGGGATGATCACTCTCACCCACAGTTGATTAAGTTGAC 240
Db 181 TCTGCAAGCAATTCCTCCCAAGGGGATGATCACTCTCACCCACAGTTGATTAAGTTGAC 240
QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
Db 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
QY 301 AGCATTTCAGTTAGAACAAAAAGACACCAAGCAAGAAATCTTGACCTATATATAATAA 360
Db 301 AGCATTTCAGTTAGAACAAAAAGACACCAAGCAAGAAATCTTGACCTATATATAATAA 360
QY 361 GGTGTACATGCTAATGGGAATATGGAATGACAGACGACCTCAAAATCTATNGTTAA 420
Db 361 GGTGTACATGCTAATGGGAATATGGAATGACAGACGACCTCAAAATCTATNGTTAA 420
QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCTTGTGCTGGAATGCTTCAGGACCC 480
Db 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCTTGTGCTGGAATGCTTCAGGACCC 480
QY 481 AAACCAATATGACCCCTATTTACATCCAGAAAGACCCAGACCCGGAATCTTGCTT 540
Db 481 AAACCAATATGACCCCTATTTACATCCAGAAAGACCCAGACCCGGAATCTTGCTT 540
QY 541 ATCTGGAATGAAAAATGAAGGCTACATCTCTGCGAAGACAGATGAGAAAGAGCAATAC 600
Db 541 ATCTGGAATGAAAAATGAAGGCTACATCTCTGCGAAGACAGATGAGAAAGAGCAATAC 600
QY 601 ACCAATTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTCATGGA 660
Db 601 ACCAATTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTCATGGA 660
QY 661 TAAATTACCTCAGGAAGTCAATCAATCAAGTTGAAGAAAGACGCTTAATCACTAC 720
Db 661 TAAATTACCTCAGGAAGTCAATCAATCAAGTTGAAGAAAGACGCTTAATCACTAC 720
QY 721 AACGGGATGATGCTACAAATGTAAGACCAAGAGCTCAAAAACATCTGTTGGGATAT 780
Db 721 AACGGGATGATGCTACAAATGTAAGACCAAGAGCTCAAAAACATCTGTTGGGATAT 780
QY 781 TTACAATACAGACGAATACGTTGGCTATCCAGACGATGAATTTGCAAGTCTTACCAT 840
Db 781 TTACAATACAGACGAATACGTTGGCTATCCAGACGATGAATTTGCAAGTCTTACCAT 840
QY 841 TGTGTGATGTTTCTAAGCGTAAGTCAATTTGCCAGCTAGGAGACGCAATCAAGTAA 900
Db 841 TGTGTGATGTTTCTAAGCGTAAGTCAATTTGCCAGCTAGGAGACGCAATCAAGTAA 900
QY 901 TGTTCCTTGGGAATTAACCAAGAGATGAGAAACAAACCGGACATGGGGATTAATAAGAA 960
Db 901 TGTTCCTTGGGAATTAACCAAGAGATGAGAAACAAACCGGACATGGGGATTAATAAGAA 960
QY 961 ACCGATCAGACATATGCTCTGCTTGGAGTAGGTCTACAGATTCAACTGCTACTAT 1020
Db 961 ACCGATCAGACATATGCTCTGCTTGGAGTAGGTCTACAGATTCAACTGCTACTAT 1020
QY 1021 CGTTACAGATGAGCCCTATTAACCTCTGGGACAAATACTCTTTATTAACGGGATAG 1080
Db 1021 CGTTACAGATGAGCCCTATTAACCTCTGGGACAAATACTCTTTATTAACGGGATAG 1080
```

```
Db 1021 CGTTACAGATGAGCCCTATTAACCTCTGGGACAAATACTCTTTATTAACGGGATAG 1080
QY 1081 GGGCTACTTTGGCAACATATCACCCTTGCAATACGCCCTGCAACATATGGGAAAGCTCCAGC 1140
Db 1081 GGGCTACTTTGGCAACATATCACCCTTGCAATACGCCCTGCAACATATGGGAAAGCTCCAGC 1140
QY 1141 CGTGGAAACTCTAAACAAAGTGGAGCTCAACCGGCCCAAGACTTTCGTAATATGGTCTAGC 1200
Db 1141 CGTGGAAACTCTAAACAAAGTGGAGCTCAACCGGCCCAAGACTTTCGTAATATGGTCTAGC 1200
QY 1201 AATGAGTACCCCAAGTATTTCACTACTCAATATGCCATTTCAAGTAACACACCAATCAGA 1260
Db 1201 AATGAGTACCCCAAGTATTTCACTACTCAATATGCCATTTCAAGTAACACACCAATCAGA 1260
QY 1261 CAAAAAATATGAGACCAAGTATGAAAAAGATGGCTGCTTACGCTCTTTCGCAAAATGG 1320
Db 1261 CAAAAAATATGAGACCAAGTATGAAAAAGATGGCTGCTTACGCTCTTTCGCAAAATGG 1320
QY 1321 TGAACCTTACTATTAACCAATGATATTCATTAAGTGGCTTTAGTATGGGAGTGAATA 1380
Db 1321 TGAACCTTACTATTAACCAATGATATTCATTAAGTGGCTTTAGTATGGGAGTGAATA 1380
QY 1381 AGATGTTCTATATGTCGGAACCTGTCGCATGGAAGAAACGACACCTATATGATAGCCGA 1440
Db 1381 AGATGTTCTATATGTCGGAACCTGTCGCATGGAAGAAACGACACCTATATGATAGCCGA 1440
QY 1441 CATGATGAAAACTCTTGAATATGGAACCTGAGACGAGAAATCTATCTTGGCTGCC 1500
Db 1441 CATGATGAAAACTCTTGAATATGGAACCTGAGACGAGAAATCTATCTTGGCTGCC 1500
QY 1501 TCAGGCTGTGTAAACACAGAACCTCTAATATACAGAGAGAAATGAAAAACCATCAA 1560
Db 1501 TCAGGCTGTGTAAACACAGAACCTCTAATATACAGAGAGAAATGAAAAACCATCAA 1560
QY 1561 GACCTCTCAATTTGTAGCAGCTGATGAACATTTTGGTGGCTATACGGTAAATTTTCAT 1620
Db 1561 GACCTCTCAATTTGTAGCAGCTGATGAACATTTTGGTGGCTATACGGTAAATTTTCAT 1620
QY 1621 GGGTGTATGAGCAGGCTATTTCTACCGTGTGACACCACTTGTAGCAATGGCCTTACGGT 1680
Db 1621 GGGTGTATGAGCAGGCTATTTCTACCGTGTGACACCACTTGTAGCAATGGCCTTACGGT 1680
QY 1681 CGGTGCCAAAGTTTACCGCTCTATATGATGACCTTACCTGCTGGAAGAAATCCAGAGA 1740
Db 1681 CGGTGCCAAAGTTTACCGCTCTATATGATGACCTTACCTGCTGGAAGAAATCCAGAGA 1740
QY 1741 TTGGAATATATACAGAGGGGCTCTACAGAAATGAGGAATTTGTAATTAATATGGTCTCG 1800
Db 1741 TTGGAATATATACAGAGGGGCTCTACAGAAATGAGGAATTTGTAATTAATATGGTCTCG 1800
QY 1801 TTTACGTGGAACCTCACCTGCTCCACAACACCCCAATCAACTGAAGTTCAAGCTCATC 1860
Db 1801 TTTACGTGGAACCTCACCTGCTCCACAACACCCCAATCAACTGAAGTTCAAGCTCATC 1860
QY 1861 ATCAGATAGTTCAACTTACAGAGTCTAGCTCAACACATCCCAAGCAATAATATAGTACGAC 1920
Db 1861 ATCAGATAGTTCAACTTACAGAGTCTAGCTCAACACATCCCAAGCAATAATATAGTACGAC 1920
QY 1921 TACCAATCTTAACATATATACGCAACATCAATATCAACCCCTATATCAACAAATTCAGAA 1980
Db 1921 TACCAATCTTAACATATATACGCAACATCAATATCAACCCCTATATCAACAAATTCAGAA 1980
QY 1981 TCCCTCAACGACACACCA 1999
Db 1981 TCCCTCAACGACACACCA 1999
```

RESULT 2
BD003825/c 10711 bp DNA linear PAT 31-JAN-2002
LOCUS BD003825
DEFINITION Polynucleotide of Streptococcus pneumoniae and sequence.
ACCESSION BD003825
VERSION BD003825.1 GI:18631786

KEYWORDS	JP 2001501833-A/145.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 10711)
AUTHORS	Kunsch,C.A., Choi,G.H., Dillon,P.J., Rosen,C.A., Bara,S.C., Fannon,M. and Dougherty,B.A.
TITLE	Polynucleotide of Streptococcus pneumoniae and sequence
JOURNAL	Patent: JP 2001501833-A 145 13-FEB-2001.
COMMENT	HUMAN GENOME SCIENCES INC
OS	Unidentified
PN	JP 2001501833-A/145
PD	13-FEB-2001
PF	30-OCT-1997 JP 1998520718
PR	31-OCT-1996 US 60/029960
PI	CHARLES A KUNSCH,GIL H CHOI,PATRICK J DILLON,CRAIG A ROSEN,PI STEVEN C BARASH,
PI	MICHAEL FANNON,BRIAN A DOUGHERTY
PC	C12N15/09,A01K67/027,C07K14/315,C07K16/12,C12N1/15,C12N1/19,
PC	C12N1/21,
PC	C12N15/10,C12P21/02,C12Q1/68,G06F17/30,C12N15/00,C12N5/00,PC
G06F15/40	
CC	Strandedness: Double;
CC	Topology: Linear;
FH	Key
FT	source
FT	1..10711
FT	/organism='Unidentified'.
FEATURES	
source	location/Qualifiers
	1..10711
	/organism='unidentified'
	/db_xref='taxon:32644'
BASE COUNT	3280 a 2056 c 2412 g 2963 t
ORIGIN	
Query Match	100.0%; Score 1999; DB 6; Length 10711;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1999: Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	TAAATTCACGCATTAATAATCACTCATCTGTCGACTTGGTCTGAAAGCGCGGTCAA 60
9767	TAAATTCACGCATTAATAATCACTCATCTGTCGACTTGGTCTGAAAGCGCGGTCAA 9708
61	TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGCAATCGTTTCATCGAAGACA 120
9707	TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGCAATCGTTTCATCGAAGACA 9648
121	TGCGTTCTTCGACACACAGGGGGATGATNCATCCGTATCCTGSGAGCTTTCTTGCGCAA 180
9647	TGCGTTCTTCGACACACAGGGGGATGATNCATCCGTATCCTGSGAGCTTTCTTGCGCAA 9588
181	TCTGGAAGCAATCCCTCCACAGGGGATGATNCATCTCCACCAAGTTGATTAAGTGC 240
9587	TCTGGAAGCAATCCCTCCACAGGGGATGATNCATCTCCACCAAGTTGATTAAGTGC 9528
241	TTACTTTTCAACTTCGACTTCGACACAGCTATTTCTGTAAGGCTCAAGAGCTTGGTT 300
9527	TTACTTTTCAACTTCGACTTCGACACAGCTATTTCTGTAAGGCTCAAGAGCTTGGTT 9468
301	AGCGATTAGTTAGACAAAAGAACCAAGCAAGAAATCTTGACCTACTATATATAATTA 360
9467	AGCGATTAGTTAGACAAAAGAACCAAGCAAGAAATCTTGACCTACTATATATAATTA 9408
361	GSTTCACATGCTATATGGGAATATGGAATGACAGACAGCTCAAAACTACTATAGTTAA 420
9407	GSTTCACATGCTATATGGGAATATGGAATGACAGACAGCTCAAAACTACTATAGTTAA 9348
421	AGACCTCAATAATTTAAAGTTTACTCTCAGTTAGCTTGCTGGGTGGAATGCTCCAGGACC 480
9347	AGACCTCAATAATTTAAAGTTTACTCTCAGTTAGCTTGCTGGGTGGAATGCTCCAGGACC 9288
481	AAACCAATATGACCCCTATTCACATCCAGAGACAGCCCAAGACGCGCGAACTTGCTTT 540

Dd	9287	AAACCAATATGACCCCTTATTCACATCCAGACAGCCCAACGCCCAACTTGCTT	9228
Qy	541	ATTCGAATGAAAAATACAGGCTACATCTCTGCTGAACGTTATGAAAAAGCATCAATAC	600
Dd	9227	ATTCGAATGAAAAATACAGGCTACATCTCTGCTGAACGTTATGAAAAAGCATCAATAC	9168
Qy	601	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAGTAATTACCTGCTTACATGGA	660
Dd	9167	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAGTAATTACCTGCTTACATGGA	9108
Qy	661	TAAATTACCTCAAGGAAGTCAATCAATCAAGTTGGAAGAAAGAACAGGCTAATACCTCAC	720
Dd	9107	TAAATTACCTCAAGGAAGTCAATCAATCAAGTTGGAAGAAAGAACAGGCTAATACCTCAC	9048
Qy	721	AACGGGATGGATGTTCTACACAATGATAGACACAGAAGCTCAAAAACATCTGGGATAT	780
Dd	9047	AACGGGATGGATGTTCTACACAATGATAGACACAGAAGCTCAAAAACATCTGGGATAT	8988
Qy	781	TTCAATACAGACGAATAGCTTGCTATTCACAGATGAATTGCAAGTGGCTTACCAT	840
Dd	8987	TTCAATACAGACGAATAGCTTGCTATTCACAGATGAATTGCAAGTGGCTTACCAT	8928
Qy	841	TGTTGATGTTTCTACAGGTAAGTCANTGGCCAGCTAGAGACACGCCATCACTCAAGTA	900
Dd	8927	TGTTGATGTTTCTACAGGTAAGTCANTGGCCAGCTAGAGACACGCCATCACTCAAGTA	8868
Qy	901	TGTTTCCCTGGGATTTAACCAAGACATAGAAACAAACCCGACTGGGATCTACATATA	960
Dd	8867	TGTTTCCCTGGGATTTAACCAAGACATAGAAACAAACCCGACTGGGATCTACATATA	8808
Qy	961	ACCGATCAAGACTATGCTCTGCTGCTTGAAGTACGGTGTACGATTCAACTGCTACTAT	1020
Dd	8807	ACCGATCAAGACTATGCTCTGCTGCTTGAAGTACGGTGTCTACGATTCAACTGCTACTAT	8748
Qy	1021	CGTTACGATGAGCCCTTAATCACTACCTGGGACAAATACTCTGTTTATAACTGGGATAG	1080
Dd	8747	CGTTACGATGAGCCCTTAATCACTACCTGGGACAAATACTCTGTTTATAACTGGGATAG	8688
Qy	1081	GGGCTACTTGGCAACATCACTTGGCAATCGCCCTGCAACATCGCCAAAGCTGCCAGC	1140
Dd	8687	GGGCTACTTGGCAACATCACTTGGCAATCGCCCTGCAACATCGCCAAAGCTGCCAGC	8628
Qy	1141	CGTGGAAATCTTAAACAAGGTGGGACTCAACCGCGCAAGACTTTCCTAATGATCTAGG	1200
Dd	8627	CGTGGAAATCTTAAACAAGGTGGGACTCAACCGCGCAAGACTTTCCTAATGATCTAGG	8568
Qy	1201	AATGACTACCCAGATTTATCACTACCTCAATGCCATTTAAGTAAACACACCGAATCGA	1260
Dd	8567	AATGACTACCCAGATTTATCACTACCTCAATGCCATTTAAGTAAACACACCGAATCGA	8508
Qy	1261	CAAAAAATATGAGACCAAGTGTGAAAAAGATGGTGTGTGTCAGCGCTTGGCAATAGG	1320
Dd	8507	CAAAAAATATGAGACCAAGTGTGAAAAAGATGGTGTGTGTCAGCGCTTGGCAATAGG	8448
Qy	1321	TGGAATCTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTATGGAGTGAATA	1380
Dd	8447	TGGAATCTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTATGGAGTGAATA	8388
Qy	1381	AGAGTCTCTAATGTGGGAATCGTGCCATGAAGAAGACAGCCATATGATGACCGA	1440
Dd	8387	AGAGTCTCTAATGTGGGAATCGTGCCATGAAGAAGACAGCCATATGATGACCGA	8328
Qy	1441	CATATGAAAAAGTCTTGACTTATGAGATGAGCAATGCTATCTGCTGGCTGCC	1500
Dd	8327	CATATGAAAAAGTCTTGACTTATGAGATGAGCAATGCTATCTGCTGGCTGCC	8268
Qy	1501	TCAGAGCTGTAAAAACAGAACCTCTAATATACAGACGAGAAATGAAAAACCAATCAA	1560
Dd	8267	TCAGAGCTGTAAAAACAGAACCTCTAATATACAGACGAGAAATGAAAAACCAATCAA	8208
Qy	1561	GACCTCTCAATTTGTAGCACCTGATGATCAATTTGCTGGCTATACGGCTAATATTCAAT	1620
Dd	8207	GACCTCTCAATTTGTAGCACCTGATGATCAATTTGCTGGCTATACGGCTAATATTCAAT	8148

Oy	1621	GGCTGTATGACAGAGCTATTCTAACCGTCTGACACCACTTGAGGCAATGGCCTTACGCT	1680	gene	/db_xref="taxon:170187"
					complement(114..2273)
					/gene="SP0369"
Db	8147	GGCTGTATGACAGAGCTATTCTAACCGTCTGACACCACTTGAGGCAATGGCCTTACGCT	8088	CDS	complement(114..2273)
					/gene="SP0369"
Oy	1681	CGCTGCCAAGTTTACCGCTCTATGATGACCTACTGCTCTGAAGAGCAATCCAGAGA	1740	gene	/transl_table=1
					/codon_start=1
Db	8087	CGCTGCCAAGTTTACCGCTCTATGATGACCTACTGCTCTGAAGAGCAATCCAGAGA	8028	CDS	/product="penicillin-binding protein 1A"
					/protein_id="AAK74536.1"
Oy	1741	TTTGAATATACCAAGGGGCTCTACAGAAATGGAATTTTCATTTTAAATATGGTCTCG	1800	gene	/db_xref="GI:14971838"
					/transl_table=1
Db	8027	TTTGAATATACCAAGGGGCTCTACAGAAATGGAATTTTCATTTTAAATATGGTCTCG	7968	CDS	/product="penicillin-binding protein 1A"
					/protein_id="AAK74536.1"
Oy	1801	TTTCTGCGGAAGTACCTGCTCCACAAACCCCATCACTGAAGTTCAAGCTCATC	1860	gene	/transl_table=1
					/codon_start=1
Db	7967	TTTCTGCGGAAGTACCTGCTCCACAAACCCCATCACTGAAGTTCAAGCTCATC	7908	CDS	/product="penicillin-binding protein 1A"
					/protein_id="AAK74536.1"
Oy	1861	ATCAGATAGTTCACCTTACAGCTAGCTCAACCACTCCACAGCAAAATATATAGTACGAC	1920	gene	/transl_table=1
					/codon_start=1
Db	7907	ATCAGATAGTTCACCTTACAGCTAGCTCAACCACTCCACAGCAAAATATATAGTACGAC	7848	CDS	/product="penicillin-binding protein 1A"
					/protein_id="AAK74536.1"
Oy	1921	TACCAATCTCTAATATATAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	1980	gene	/transl_table=1
					/codon_start=1
Db	7847	TACCAATCTCTAATATATAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	7788	CDS	/product="penicillin-binding protein 1A"
					/protein_id="AAK74536.1"
Oy	1981	TGCTCAACCAAGCAACCA 1999		gene	/transl_table=1
					/codon_start=1
Db	7787	TGCTCAACCAAGCAACCA 7769		CDS	/product="penicillin-binding protein 1A"
					/protein_id="AAK74536.1"
RESULT 3					
AE007349/c					
LOCUS	AE007349	11334 bp	DNA	linear	BCT 31-AUG-2001
DEFINITION	Streptococcus pneumoniae TIGR4 section 32 of 194 of the complete genome.				
ACCESSION	AE007349	AE005672			
VERSION	AE007349.1	GI:14971837			
KEYWORDS	Streptococcus pneumoniae TIGR4.				
SOURCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.				
ORGANISM	Streptococcus pneumoniae TIGR4.				
REFERENCE	1 (bases 1 to 11334)				
AUTHORS	Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Salzberg, S.L., Lewis, M.R., Radune, D., Holtzapple, E., Khouli, H., Wolf, A.M., Uterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Dickinson, T., Hickey, E.K., Holt, I.E., Loftus, B.J., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.				
TITLE	Complete genome sequence of a virulent isolate of Streptococcus pneumoniae				
JOURNAL	Science. 293 (5529), 498-506 (2001)				
MEDLINE	21357209				
PUBMED	11463916				
REFERENCE	2 (bases 1 to 11334)				
AUTHORS	Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Lewis, M.R., Radune, D., Holtzapple, E., Khouli, H., Wolf, A.M., Uterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Gesuwan, P., Hickey, E.K., Holt, I.E., Loftus, B.J., Ujwal, M.L., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
FEATURES	Location/Qualifiers				
Source	1. 11334				
	/organism="Streptococcus pneumoniae TIGR4"				
	/strain="TIGR4"				

gene
CDS
/translation="MKKEFNLIATVAAGLEAVVGREVELGYDCOVENGRVRFQGDVRA
AIETINLMLRADRIKIIIVGFPAPKTEPEELFOGVFALDMENYILPLGARFPTSKCYK
SKIHNEPSVALISKKAIVVKKLOKHARBEGVPLMENGPEEKIEVSLIKDVTATIDIT
GSSLFRKRGYRTEKGAPIKENMAAIIQLSMWYDKPLIDPTCGSGTFCIYAVIARK
MAPGLRSPAFEEKNWISDRLIOVRTBAKVRLELMDICDIDARVEIARANA
OVAGVADITFFKQVODLRSDKINGVILSNPPRGESDAGVTKLYAEMGOVFAPL
KTMSEFLITSEDAFEESKYSQADKKRLYNGLTKVDLYQYFGQKVRQEVK"
5514..6908
/gene="Sp0374"
5514..6908
/gene="Sp0374"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK74541.1"
/db_xref="GI:14971843"
/translation="MSKKRRNRHKEGOEPDFDEAKELTVGOAIRKNEEVEGVLP
EDSLIDRYVKOHDEIEADKFAATROYKKEEVEVQSDDLIOEREAVERKESEASEV
PSSDILPLIDDEBOGLDPLILDENPTMEFEVEEONLSLDOESKRRKGF
ILTYLAIVSYICISAYYVROVARSTKEIFTSSTTANOSDVDFNTLYDAFTDSN
KALKNSQFPLSOLTKLIDKLBGSRHTLAKSKYDSLATQIKAIODVNAQFEKPAIV
DGYLDITNAKAKSDAKFTDIKGTNELDKVLDAKISLKSOOTSTSSSSSQTSSSSS
QASNTTSEKPSSESTRSRSEVNMGLSAGVAVORSARVAYNOSAIIDSNNSSAW
DFADGVLEOILATSRSRGYITGDOYLEERVIYVNGVYLNKPDGYTLFTLNCKTGY
FVGNAGHADDLDY"
6984..8408
/gene="Sp0375"
6984..8408
/gene="Sp0375"
/note="identified by match to TIGR protein family HMM
TIGR00873"
/codon_start=1
/transl_table=11
/product="6-phosphogluconate dehydrogenase,
decarboxylating"
/protein_id="AAK74542.1"
/db_xref="GI:14971844"
/translation="MTKANFCVGMAYMGRNLAINISRGYTAIVNRSKKTEEDVIA
CHEKPNVPSYDVESEYVNSIEKPRIMLMVOAGGTDTTIALPLHLDKGLIDGNG
TFKKDITRNEELANSQINFITGVSGSEGALGPISMPGQKEAEVLVADVLEIS
AKAPEDKPCVYIIGPDAGHYVMVHNGIEYDMQILAESYDLMOHLIGLSADMAE
IFTEPMNGELSDSYLIEITADILSRKDEGDGDPVVDYILDAAGKGTGKSSQSLD
GVPLSTIESEVFAVYISTYKEERYHASKVLPRPAFMEGRKAEILIKIRATFESIT
ISYNOGPAQLRASKENNNMLPRADIASIMPDGCTIISRFLOKTTDYNRRDANLIL
LDEYFLDYATAKYQADVADIVALAQAQVPPTEFAATYFEDSYRSADLPALVIAQRD
YFGAHTYQRKDKEGTFHYSWDEK"
8420..9109
/gene="Sp0376"
8420..9109
/gene="Sp0376"
/note="Sp0376"
/codon_start=1
/transl_table=11
/product="DNA-binding response regulator"
/protein_id="AAK74543.1"
/db_xref="GI:14971845"
/translation="MGRKILLEKERNLAHFLSELOKEQYRVDLVERGOKALSMAIO
TDYDMLNVLNGLDMAODFAEKLSRTKPSVIMILDHWELOLELVEYVQFAVSILY
KPVLEMLVARISAIIFRGDRFTIDHCSLMKVPRIYRMLRIDVHHHTYRGEMALTR
REYDLATLMGSKKVLTRQELLESVMKYESATEINIVDYIRYRSKLDYVGQASYIK
TVRGVYTMQE"
9208..10230
/gene="Sp0377"
9208..10230
/gene="Sp0377"
/note="Sp0377"
/codon_start=1
/transl_table=11
/product="choline binding protein C"
/protein_id="AAK74544.1"
/db_xref="GI:14971846"
/translation="MKLLKRMVOYALAVFFEGGLIATNTVFANTTGRFRVDKNNRYV
KDHKAIVMHRIDCKTYFGDIGEMVGMOLYLEIPGQYRNDLPDNPVNIIGQOEK
YFGQDGLALLEQTDKQVLEAKTSENTGVYGEQYPLSAEKRTYFEDNNYAKTGMWIE

EGHWYLLNKLGFGDDSYNPLPIGEVAKGMTDFHVTIIDRSKPAPWYLDASGKML
TDMQVKNKRYVFCSSGSMATGMVVRKRYVILUNKNDMDTGMQYLOKMKYILRSSG
AMTGQVODSSTWYIYLDPSNDMKIIGWKVANGKMYILNSNAAMTGSOTIDGKYVNFPA
SSGERT"
10248..11246
/gene="Sp0378"
10248..11246
/gene="Sp0378"
/codon_start=1
/transl_table=11
/product="choline binding protein J"
/protein_id="AAK74545.1"
/db_xref="GI:14971847"
Query Match 100.0%; Score 1999; DB 1; Length 11334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TAAATCTACGACATTAATAATCACTCATTTGCTGACCTGGTTCGAAGCGCGTCAA 60
|||||
Db 2115 TAAATCTACGACATTAATAATCACTCATTTGCTGACCTGGTTCGAAGCGCGTCAA 2056
61 TGCCCAAGCTAATGATATTCGCCAGATTTGGTTAAGGCAATCGTTCTATCGAAGACCA 120
|||||
Db 2055 TGCCCAAGCTAATGATATTCGCCAGATTTGGTTAAGGCAATCGTTCTATCGAAGACCA 1996
121 TCGCTTTTCGACACACAGGGGATTTGATACCATTCGGTATCTGGGACGTTTTCGCCAA 180
|||||
Db 1995 TCGCTTTTCGACACACAGGGGATTTGATACCATTCGGTATCTGGGACGTTTTCGCCAA 1936
181 TGTGCAAGCAATTCCTCCCAAGGTGATGATCACTGCAACCCCAAGTTGATTAAGTTGAC 240
|||||
Db 1935 TGTGCAAGCAATTCCTCCCAAGGTGATGATCACTGCAACCCCAAGTTGATTAAGTTGAC 1876
241 TTACTTTTCAACTTGACACTTCGACACAGTATTTCTGTAAAGGCTCAGAAAGCTTGGTT 300
|||||
Db 1875 TTACTTTTCAACTTGACACTTCGACACAGTATTTCTGTAAAGGCTCAGAAAGCTTGGTT 1816
301 AGCGATTACGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATTAATTA 360
|||||
Db 1815 AGCGATTACGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATTAATTA 1756
361 GGTTCACATGCTAATGGGAACCTATGGAATGAGACAGACAGAGCTCAAACTACTATAGTTAA 420
|||||
Db 1755 GGTTCACATGCTAATGGGAACCTATGGAATGAGACAGACAGAGCTCAAACTACTATAGTTAA 1696
421 AGACCTCAATATTTAAGTTTACTCTAGTTTACCTTGTGCTGGATGGAATGCTTCAGACACC 480
|||||
Db 1695 AGACCTCAATATTTAAGTTTACTCTAGTTTACCTTGTGCTGGATGGAATGCTTCAGACACC 1636
481 AAACCAATATGACCCCTATTTACATTCAGACAGACAGACCCCAAGACCGCGAACTTGCTTT 540
|||||
Db 1635 AAACCAATATGACCCCTATTTACATTCAGACAGACAGACCCCAAGACCGCGAACTTGCTTT 1576
541 ATCTGAATATGAAAATCAAGGCTACATCTCTGCTGAACGTTGAGAAAGCGTGAATTAAC 600
|||||
Db 1575 ATCTGAATATGAAAATCAAGGCTACATCTCTGCTGAACGTTGAGAAAGCGTGAATTAAC 1516
601 ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGCAAGTATTAACCTGCTTACATGGA 660
|||||
Db 1515 ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGCAAGTATTAACCTGCTTACATGGA 1456
661 TAAATTAAGTCAAGGAAGTCAATCAAGTTGAGAGAGAAAGACGCTTAACCTACTAC 720
|||||
Db 1455 TAAATTAAGTCAAGGAAGTCAATCAAGTTGAGAGAGAAAGACGCTTAACCTACTAC 1396
721 AACCTGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
|||||
Db 1395 AACCTGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 1336
781 TTACATATCAGACGAATACGTTGCTATTCAGACAGATGAATTCGAAGTGCCTTACACAT 840
|||||
Db 1335 TTACATATCAGACGAATACGTTGCTATTCAGACAGATGAATTCGAAGTGCCTTACACAT 1276

QY 841 TGTGATGTTTCTAAGGTAAGTCAATGCGCCAGCTAGAGCAGCCATCAGTCAAGTAA 900
 |||||
 Db 1275 TGTGATGTTTCTAAGGTAAGTCAATGCGCCAGCTAGAGCAGCCATCAGTCAAGTAA 1216
 |||||
 QY 901 TGTTCCTTGGGAATTAACGACAGTAAACAAACCGGACATGGGATCACTATGAA 960
 |||||
 Db 1215 TGTTCCTTGGGAATTAACGACAGTAAACAAACCGGACATGGGATCACTATGAA 1156
 |||||
 QY 961 ACCGATCAGACACTATGCTCTGCTTGGAGTACGCTGTCACGATTCACCTGCTACTAT 1020
 |||||
 Db 1155 ACCGATCAGACACTATGCTCTGCTTGGAGTACGCTGTCACGATTCACCTGCTACTAT 1096
 |||||
 QY 1021 CGTTACGATGAGCCCTAATACTACCTGGGACAAATCTCCTTTTATTAACGGGATAG 1080
 |||||
 Db 1095 CGTTACGATGAGCCCTAATACTACCTGGGACAAATCTCCTTTTATTAACGGGATAG 1036
 |||||
 QY 1081 GGGCTACTTTGGCAACATCAGCTTGCATACGCTTGCACAAATCGGAAACGTCGCCAGC 1140
 |||||
 Db 1035 GGGCTACTTTGGCAACATCAGCTTGCATACGCTTGCACAAATCGGAAACGTCGCCAGC 976
 |||||
 QY 1141 CGTGAACCTGTAACGAGTGGAGTCAACCGGCGCAAGCTTCTTAATGGCTAGG 1200
 |||||
 Db 975 CGTGAACCTGTAACGAGTGGAGTCAACCGGCGCAAGCTTCTTAATGGCTAGG 916
 |||||
 QY 1201 AATGACTACCCAGATATTCCTACTCTCAATGCGCATTTCAAGTAAACACGGAATCAGA 1260
 |||||
 Db 915 AATGACTACCCAGATATTCCTACTCTCAATGCGCATTTCAAGTAAACACGGAATCAGA 856
 |||||
 QY 1261 CAAAAAATATGAGCAGTATGTAAGAGATGGCTGCTTACGCTGCTTTGCAATGG 1320
 |||||
 Db 855 CAAAAAATATGAGCAGTATGTAAGAGATGGCTGCTTACGCTGCTTTGCAAAAGG 796
 |||||
 QY 1321 TGAACCTACTATTAACCAATGTAATTCACATTAAGTCTCTTTAGTATGGAGTGAAGA 1380
 |||||
 Db 795 TGAACCTACTATTAACCAATGTAATTCACATTAAGTCTCTTTAGTATGGAGTGAAGA 736
 |||||
 QY 1381 AGAGTCTCTAATGTGCGAATCTGCCATGAGAGAAACGACACCTATATGATGACCGA 1440
 |||||
 Db 735 AGAGTCTCTAATGTGCGAATCTGCCATGAGAGAAACGACACCTATATGATGACCGA 676
 |||||
 QY 1441 CATATGTAACCAATGCTGATGTAAGAGTGAACGAGAAATGCGTATCTGCTGGCTCCC 1500
 |||||
 Db 675 CATATGTAACCAATGCTGATGTAAGAGTGAACGAGAAATGCGTATCTGCTGGCTCCC 616
 |||||
 QY 1501 TCAGGCTGTGTAAACAGAGAACTCTAATATACAGAGAGAAATGGAACCAATCA 1560
 |||||
 Db 615 TCAGGCTGTGTAAACAGAGAACTCTAATATACAGAGAGAAATGGAACCAATCA 556
 |||||
 QY 1561 GACCTCTCAATTTGTAGCAGCTGATGAATATTTGCTGGCTATACGGGTAAATATTCAT 1620
 |||||
 Db 555 GACCTCTCAATTTGTAGCAGCTGATGAATATTTGCTGGCTATACGGGTAAATATTCAT 496
 |||||
 QY 1621 GGGCTGTATGACAGGCTATTTAAACGCTGACACCACTTGTAGCAATGGCTTACGGT 1680
 |||||
 Db 495 GGGCTGTATGACAGGCTATTTAAACGCTGACACCACTTGTAGCAATGGCTTACGGT 436
 |||||
 QY 1681 CGGTGCAAAAGTTTACGGCTATATGATGAGCACTGCTGTAAGGAAGCAATCCAGAGA 1740
 |||||
 Db 435 CGGTGCAAAAGTTTACGGCTATATGATGAGCACTGCTGTAAGGAAGCAATCCAGAGA 376
 |||||
 QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAATTTCTATTTAAATGGTGTCCG 1800
 |||||
 Db 375 TTGGAATATACAGAGGGGCTCTACAGAAATGGAATTTCTATTTAAATGGTGTCCG 316
 |||||
 QY 1801 TTCTAGCTGGAATCTGCTGTCACAAACACCCCATCAACTGAAGTTCAAGCTATC 1860
 |||||
 Db 315 TTCTAGCTGGAATCTGCTGTCACAAACACCCCATCAACTGAAGTTCAAGCTATC 256
 |||||
 QY 1861 ATCAGATAGTCACTTACAGCTAGCTACAGCAACGCAACCAATATATATAGTACGAC 1920
 |||||
 Db 255 ATCAGATAGTCACTTACAGCTAGCTACAGCAACGCAACCAATATATATAGTACGAC 196
 |||||

QY 1921 TACCAATCTTACATATATACGACAAATCAATCAACCCCTGATCAACAAATTCAGAA 1980
 |||||
 Db 195 TACCAATCTTACATATATACGACAAATCAATCAACCCCTGATCAACAAATTCAGAA 136
 |||||
 QY 1981 TCTTCAACCCAGACACCA 1999
 |||||
 Db 135 TCTTCAACCCAGACACCA 117
 |||||

RESULT 4
 AF210745
 LOCUS 2160 bp DNA linear BCF 17-SEP-2001
 DEFINITION Streptococcus pneumoniae isolate SP 1261 penicillin-binding protein
 1A (pbp1a) gene, complete cds.
 ACCESSION AF210745
 VERSION AF210745.1 GI:6563336
 KEYWORDS
 SOURCE Streptococcus pneumoniae.
 ORGANISM Streptococcus pneumoniae.
 Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 Streptococcus.
 REFERENCE 1 (bases 1 to 2160)
 AUTHORS Fertoni, A. and Berche, P.
 TITLE Alterations to penicillin-binding proteins 1a, 2b and 2x amongst
 penicillin-resistant clinical isolates of Streptococcus pneumoniae
 serotype 23f from the nasopharyngeal flora of children
 J. Med. Microbiol. 50 (9), 828-832 (2001)
 MEDLINE 21432820
 PUBMED 11549185
 REFERENCE 2 (bases 1 to 2160)
 AUTHORS Fertoni, A. and Berche, P.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-1999) Microbiology, Hospital
 Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
 FEATURES
 source
 1..2160
 /organism="Streptococcus pneumoniae"
 /serotype="23f"
 /isolate="SP 1261"
 /db_xref="taxon:1313"
 1..2160
 /gene="pbp1a"
 1..2160
 /gene="pbp1a"
 /transl_start=1
 /transl_table=11
 /product="penicillin-binding protein 1a"
 /protein_id="AFI17255.1"
 /db_xref="GI:6563337"
 /translation="MNKPTLRILKYLISFLSLVIAIVLGGVFEYVSKAPLSLSE
 SKLVAATTSKTYDNKNQILADLSERRVNAQANDIPDLVKAIVSIEDHFPDRHID
 TIRILAFPLRNLOSNSLOGSSTLQOLIKLYSTSPDITSRKQEAVALAOLBOK
 ATKOEILTYINKVYMSNGYMGWTAQVYAGDNLNLSIPOLALLAGMPQANQYDP
 YSHPEAQRNRNLVLSBKNQGISBQIEKAVNPTITDLSLSKASNTPAYMDNLY
 KEVINQVEETGYNLLTGMVTVNDQEKHLMDIYNDEVAAPDELOVASTIV
 DVSNGRVIQDLGARHQSNSVSGFINQAVENRDMGSTMPIEDYALAEVGYDSTAT
 IYHDEPNYEGTNPYVNMDRGVEGNTTLOYLALQOSRNPAYETLKNKGLNRKPTLN
 GLCIDVPSIHYSNASISNTSESKKYGASSEKMAAAVAAANGVTYKRPYIKHYVES
 DGESEKFSNVGTRAMKETTAHYMTDMKQVLYLTGTRNAYLAMLPAQKTVRSNTYDE
 EIEINHKTISOFAVDELFGYTRKYISMAWTGYSNLTPLVGLTLVAAKVYRSMTY
 LSEGSNPEDMNIEPEGLYRNGEYFVKNGARSTWNSPAPQPPSTESSSSSDSTSSOS
 STPTSTNNSTTTNNNTQOOSNTTPOQONQPPAP"
 BASE COUNT 700 a 521 c 408 g 531 t
 ORIGIN

Query Match 99.7%; Score 1992.6; DB 1; Length 2160;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1995; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATATAATCACTATTCGACTTGGCTTCGACGCCGCTCAA 60
 |||||
 Db 159 TAAATCTAGACAAATATAATCACTATTCGACTTGGCTTCGACGCCGCTCAA 218
 |||||

QY	61	TGCCAAGCTAATGATATTTCCCAAGATTTGGTTAAGCAATCGTTTCTATCGAAGACA	120
Dp	219	TGCCCAAGCTAATGATATTTCCCAAGATTTGGTTAAGCAATCGTTTCTATCGAAGACA	278
OY	121	TGCGTCTTGAGCACAGGGGGGATATATCCATCCGTAATCCGGGAGGCTTCTTGCGCA	180
Dp	279	TGCGTCTTGAGCACAGGGGGGATATATCCATCCGTAATCCGGGAGGCTTCTTGCGCA	338
OY	181	TCGTCAAGAATTTCCCTCCCAAGGTGGATCACTCCACCACAGTTGATTAAGTTGAC	240
Dp	339	TCGTCAAGAATTTCCCTCCCAAGGTGGATCACTCCACCACAGTTGATTAAGTTGAC	398
OY	241	TTACTTTTCAACTTCGACTTCGCAGCACACTATTTCTCGTAAGGCTCAGGAAGCTTGTT	300
Dp	399	TTACTTTTCAACTTCGACTTCGCAGCACACTATTTCTCGTAAGGCTCAGGAAGCTTGTT	458
OY	301	AGGATTCAGTTTGACAAACAAACCAACCAAGAAATCTTGACCTCTATATATAATA	360
Dp	459	AGGATTCAGTTTGACAAACCAACCAACCAAGAAATCTTGACCTCTATATATAATA	518
OY	361	GGTTACATGTCTAAATGGGAATCATATGGAATGAGACAGAGCTCAAACTACTATGTAA	420
Dp	519	GGTTACATGTCTAAATGGGAATCATATGGAATGAGACAGAGCTCAAACTACTATGTAA	578
OY	421	AGACCTCAATTAATTTAAGTTTACCTCAAGTTAGGCTTGCTGGGTGAATGCTCAGGAC	480
Dp	579	AGACCTCAATTAATTTAAGTTTACCTCAAGTTAGGCTTGCTGGGTGAATGCTCAGGAC	638
OY	481	AAACCATATAGACCCCTATTACATCCAGAGACAGCCCAAGCCGCAAACTTGCTT	540
Dp	639	AAACCATATAGACCCCTATTACATCCAGAGAGGCCCAAGCCGCAAACTTGCTT	698
OY	541	ATTCGAATGAAAAATCAAGGCTACATCGCTCGTGAACGTATGGAAGAAGCTCAATAC	600
Dp	699	ATTCGAATGAAAAATCAAGGCTACATCGCTCGTGAACGTATGGAAGAAGCTCAATAC	758
OY	601	ACCAATTTACTGATGAGACTACAAAGTCTCAAAATCAGCAAAATTAATACCTGCTACATGA	660
Dp	759	ACCAATTTACTGATGAGACTACAAAGTCTCAAAATCAGCAAAATTAATACCTGCTACATGA	818
OY	661	TAAATTACCTCAAGGAAGTCATCAATCAAGTTGAGAGAAGAACAGGCTATTAACCTACTAC	720
Dp	819	TAAATTACCTCAAGGAAGTCATCAATCAAGTTGAGAGAAGAACAGGCTATTAACCTACTAC	878
OY	721	AACGGGATGGATGTCTACACCAATGTATACCAAGAAGGCTCAAAAACATCTGTGGGATAT	780
Dp	879	AACGGGATGGATGTCTACACCAATGTATACCAAGAAGGCTCAAAAACATCTGTGGGATAT	938
OY	781	TTCAATATACAGAGAAATACGTTGCTCATCCAGACGATGAATGCAAGTCGCTTCAACAT	840
Dp	939	TTCAATATACAGAGAAATACGTTGCTCATCCAGACGATGAATGCAAGTCGCTTCAACAT	998
OY	841	TGTTGATGTTTCTAAGGTTAAAGTCATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA	900
Dp	999	TGTTGATGTTTCTAAGGTTAAAGTCATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA	1058
OY	901	TGTTTCTCTGGGAATTAACCAAGCAGTAAAGAAACACCGCAGCTGGGATCAACTATGAA	960
Dp	1059	TGTTTCTCTGGGAATTAACCAAGCAGTAAAGAAACACCGCAGCTGGGATCAACTATGAA	1118
OY	961	ACCGATACAGACATAGCTCTGCTCTGGAGTACGGTGTCTACGATTAACGCTACTAT	1020
Dp	1119	ACCGATACAGACATAGCTCTGCTCTGGAGTACGGTGTCTACGATTAACGCTACTAT	1178
OY	1021	CGTTACAGATGAGCCCTTAATCACTACCTGGGACAAATACTCCTGTTTAACTGGGATAG	1080
Dp	1179	CGTTACAGATGAGCCCTTAATCACTACCTGGGACAAATACTCCTGTTTAACTGGGATAG	1238
OY	1081	GGGCTACTTTGGCAACATCACTTGCAATACGCGCTGCAACATCGCGAAAGCTCCAGC	1140
Dp	1239	GGGCTACTTTGGCAACATCACTTTGGCAATACGCGCTGCAACATCGCGAAAGCTCCAGC	1298

Accession	Organism	Gene	Length (bp)	GC Content (%)	Reference
QY 1141	<i>CSTGGAACCTCTAACAAGAGTGGACTCAACCGCGCCAAACCTTTCTTAATGGTCTACG</i>		1200	50.0	
Db 1299	<i>CGTGAACCTCTAACAAGAGTGGACTCAACCGCGCCAAACCTTTCTTAATGGTCTACG</i>		1358	50.0	
QY 1201	<i>AATCGACTCCCAAGATTCACCTCAATGGCATTTCAAGTAACCAACCGAATCAGA</i>		1260	50.0	
Db 1359	<i>AATCGACTCCCAAGATTCACCTCAATGGCATTTCAAGTAACCAACCGAATCAGA</i>		1418	50.0	
QY 1261	<i>CAAAAAATATGAGCAAGTAAGTGAAGAAATGGCTGCTCTTACGCTCCCTTTGCAAAATGG</i>		1320	50.0	
Db 1419	<i>CAAAAAATATGAGCAAGTAAGTGAAGAAATGGCTGCTCTTACGCTCCCTTTGCAAAATGG</i>		1478	50.0	
QY 1321	<i>TGGAACTTACTATTAACCAATGTATTCATTAAGTGTCTTATAGTGGAGTGAATA</i>		1380	50.0	
Db 1479	<i>TGGAACTTACTATTAACCAATGTATTCATTAAGTGTCTTATAGTGGAGTGAATA</i>		1538	50.0	
QY 1381	<i>ACAGTCTCTAATGTGGGAACCTGCGCATGAAGGAAGCAACGACCATATATGATGACCGA</i>		1440	50.0	
Db 1539	<i>ACAGTCTCTAATGTGGGAACCTGCGCATGAAGGAAGCAACGACCATATATGATGACCGA</i>		1598	50.0	
QY 1441	<i>CATGATGAAGAAACAGTCTTGACTTATGGAACCTGGAACGAATGCGTATTCGTTGGCTCC</i>		1500	50.0	
Db 1599	<i>CATGATGAAGAAACAGTCTTGACTTATGGAACCTGGAACGAATGCGTATTCGTTGGCTCC</i>		1658	50.0	
QY 1501	<i>TCAGGCTGTAAAAACAGGAACCTCTAATCTATACAGACGAGGAAATGAAAAACCATCA</i>		1560	50.0	
Db 1659	<i>TCAGGCTGTAAAAACAGGAACCTCTAATCTATACAGACGAGGAAATGAAAAACCATCA</i>		1718	50.0	
QY 1561	<i>GACCTCTCAATTTGTAGCACTGATGACATATTTGCTGGCTATACGGGTAAATTTCAAT</i>		1620	50.0	
Db 1719	<i>GACCTCTCAATTTGTAGCACTGATGACATATTTGCTGGCTATACGGGTAAATTTCAAT</i>		1778	50.0	
QY 1621	<i>GGCTGATGAGACAGGCTATTTCTAACCGTCTGACACCACTTGTAGCAATGGCTTACGGT</i>		1680	50.0	
Db 1779	<i>GGCTGATGAGACAGGCTATTTCTAACCGTCTGACACCACTTGTAGCAATGGCTTACGGT</i>		1838	50.0	
QY 1681	<i>CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTTGAAGGAAGCAATCCAGAGA</i>		1740	50.0	
Db 1839	<i>CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTTGAAGGAAGCAATCCAGAGA</i>		1898	50.0	
QY 1741	<i>TTTGAATATATACAGAGGGGCTCTACAGAAATGGAGAAATTTGTTTAAAAATGGTGGCTCG</i>		1800	50.0	
Db 1899	<i>TTTGAATATATACAGAGGGGCTCTACAGAAATGGAGAAATTTGTTTAAAAATGGTGGCTCG</i>		1958	50.0	
QY 1801	<i>TTCTACGTGGAACTCAACCTGCTCACAACAACCCCATCAACTGAAGTCAACCTCATC</i>		1860	50.0	
Db 1959	<i>TTCTACGTGGAACTCAACCTGCTCACAACAACCCCATCAACTGAAGTCAACCTCATC</i>		2018	50.0	
QY 1861	<i>ATCAGATAGTTCAACTTCAACAGTCTAGCTCAACCTCAACCTCAACGAACAATATATAGTAGAC</i>		1920	50.0	
Db 2019	<i>ATCAGATAGTTCAACTTCAACAGTCTAGCTCAACCTCAACCTCAACGAACAATATATAGTAGAC</i>		2078	50.0	
QY 1921	<i>TACCAATCTTACATATATATACGACAACTTAATACACCCCTGATATACAAATTCAGAA</i>		1980	50.0	
Db 2079	<i>TACCAATCTTACATATATATACGACAACTTAATACACCCCTGATATACAAATTCAGAA</i>		2138	50.0	
QY 1981	<i>TTCTCAACGACGACCAACCA 1999</i>				
Db 2139	<i>TTCTCAACGACGACCAACCA 2157</i>				

AUTHORS Hakenbeck, R.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.
Genetik, Ihnestr. 73, 1000 Berlin 33, FRG
REFERENCE 2 (bases 1 to 2157)
AUTHORS Martin, C., Sibold, C. and Hakenbeck, R.
TITLE Relatedness of penicillin-binding protein 1a genes from different
clones of penicillin-resistant Streptococcus pneumoniae isolated in
South Africa and Spain
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)
MEDLINE 93010977
FEATURES
Source Location/Qualifiers
1..2157
/organism="Streptococcus pneumoniae"
/strain="45607"
/db_xref="taxon:1313"
/clone="pckM1-pckM16"
1..2157
/gene="pona"
1..>2157
/gene="pona"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="CA48073.1"
/db_xref="GI:47420"
/db_xref="SMSS-PROT:004707"
translation="MNKPTILRLKYLISPLSIVIAIVLGGVEFYVSKAPSLSE
SKVATTSKRYDKNKLADLSERRRNANPDLVLAIVLSEIDHREFDRI
TIRLGLAFRLQSNLSLOGSTLTQOLKLYSTSTSDOTISKADAWALADEOK
ATKQEIILYVINKVYMSNGYIAQONVYKQDLNLSLQALALAMPAPNOYD
YSHPEADDRNLVLEEMKNGYISAEQYKAVMTPIIDGLOSLSKSNYPAYMNYL
KEVNOVEEYGNLITGMVYTNVDOEAKHLMDIYNTDEYVAYPDELOVASTY
DVSNKYVIAOLGARHOSNVEFIGNOAVETNRDQKSTMPKITDVAAPALEYGVNSTAT
IVHPEPVNRYCTNPNVNMDSRGFNITLQVALOOSRNPVAVELINKGLRATPLN
GLGTDYPSIHNSAIISSVTTSDKRYGSSSKMAAIAFAENGTYIKPMTIHWVES
DGSEKESNVGTTRAMKETTAAMTEMKMTVLVYGTGRNAVLAAPQAKTIGSNVTE
EIEHNIKTSQVAPDELFAGYTRKYSMAVWGYNRLPLVNGILTVAAKYRSMY
LSEGSNEDMNIPEGLYRNGEFVFNKNGARSTWSSPAPOQPESTESSSSSDSTSSS
STTPTSTNSTTNNNTTQSNITPTDOONOPQAPQ"

BASE COUNT 696 a 520 c 411 g 530 t
ORIGIN

Query Match 99.4%; Score 1986.2; DB 1; Length 2157;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAATAAATCAACTGCTGACTGGTGTCTGAACGCCGCGTCAA 60
DB 159 TAAATCTACGACAATAAATCAACTGCTGACTGGTGTCTGAACGCCGCGTCAA 218
QY 61 TGCCCAAGCTATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATGGAAGACCA 120
DB 219 TGCCCAAGCTATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATGGAAGACCA 278
QY 121 TCGCTTTTCGACACAGAGGGGATGATACATCCGTAATCCGAGAGCTTCTTGGCAA 180
DB 279 TCGCTTTTCGACACAGAGGGGATGATACATCCGTAATCCGAGAGCTTCTTGGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCAACCCACAGTTGATTAGTTGAC 240
DB 339 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCAACCCACAGTTGATTAGTTGAC 398
QY 241 TTACTTTCACTTCGACTTCGACACAGATTTTCTGTAAGGCTCAGGAAGCTTGGTT 300
DB 399 TTACTTTCACTTCGACTTCGACACAGATTTTCTGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCATTCAGTTAGACAAAAGCAACAGCAAGAAATCTTGACCTATATATAATA 360
DB 459 AGCATTCAGTTAGACAAAAGCAACAGCAAGAAATCTTGACCTATATATAATA 518
QY 361 GGTCTACATGTCTAATGGGAAGTATGAAAGCAGACAGCTCAAACTATATGTTAA 420

DB 519 GGTCTACATGTCTAATGGGAAGTATGAAAGCAGACAGCTCAAACTATATGTTAA 578
QY 421 AATCCTCAATATTTTAAGTTTACCTTACCTGAGTACCTTGGCTGGAAATCCCTCAGGACCC 480
DB 579 AATCCTCAATATTTTAAGTTTACCTTACCTGAGTACCTTGGCTGGAAATCCCTCAGGACCC 638
QY 481 AATCCTCAATATTTTAAGTTTACCTTACCTGAGTACCTTGGCTGGAAATCCCTCAGGACCC 540
DB 639 AATCCTCAATATTTTAAGTTTACCTTACCTGAGTACCTTGGCTGGAAATCCCTCAGGACCC 698
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGTATGAGAAAGCATCAATAC 600
DB 699 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGTATGAGAAAGCATCAATAC 758
QY 601 ACCAATTCCTGATGACATCAACAAAGTCTCAAAATCAGCAAGTATTAACCCGCTTACATGGA 660
DB 759 ACCAATTCCTGATGACATCAACAAAGTCTCAAAATCAGCAAGTATTAACCCGCTTACATGGA 818
QY 661 TAATTAACCTCAAGAAAGTATCAATCAAGTTGAAAGAAACAGGCTATTAACCTACAC 720
DB 819 TAATTAACCTCAAGAAAGTATCAATCAAGTTGAAAGAAACAGGCTATTAACCTACAC 878
QY 721 AACTGGGATGATGTCTACACAAATGTAGACCAGAGCTCAAAAACATCTGTGGATAT 780
DB 879 AACTGGGATGATGTCTACACAAATGTAGACCAGAGCTCAAAAACATCTGTGGATAT 938
QY 781 TTACAATACAGACGAATACGTTGCTTCCCTATCCAGACATGAATTGCAAGTCCCTTACCAT 840
DB 939 TTACAATACAGACGAATACGTTGCTTCCCTATCCAGACATGAATTGCAAGTCCCTTACCAT 998
QY 841 TGTATATGTTTCTAAGCTTAAAGTATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 999 TGTATATGTTTCTAAGCTTAAAGTATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1058
QY 901 TGTTCCTTGGGAATTAACCAAGAGTAAAGCAAAACCGGAGTGGGATCACTATGAA 960
DB 1059 TGTTCCTTGGGAATTAACCAAGAGTAAAGCAAAACCGGAGTGGGATCACTATGAA 1118
QY 961 ACCGATCAAGACATATGCTTCCCTGCTGAGAGTGGTCTTACAGTTCACATGCTACTAT 1020
DB 1119 ACCGATCAAGACATATGCTTCCCTGCTGAGAGTGGTCTTACAGTTCACATGCTACTAT 1178
QY 1021 GCTTCACGATGAGCCCTATTAATCACTGAGCAATATCTCTTATTAATCTGGGATAG 1080
DB 1179 GCTTCACGATGAGCCCTATTAATCACTGAGCAATATCTCTTATTAATCTGGGATAG 1238
QY 1081 GGGCTACTTTGGCAACATACCTTGAATAGCGCCCTGCAACATGCGAAAGCTCCAGC 1140
DB 1239 GGGCTACTTTGGCAACATACCTTGAATAGCGCCCTGCAACATGCGAAAGCTCCAGC 1298
QY 1141 CGTGAACCTGTAACCAAGGTGCGACTACACCGGCGCAAGCTTCCATAATGGCTAGG 1200
DB 1299 CGTGAACCTGTAACCAAGGTGCGACTACACCGGCGCAAGCTTCCATAATGGCTAGG 1358
QY 1201 AATGCACTACCAAGTATTAATCACTACTCAATGCAATTCCTCAAGTAAACACACGGAATCGA 1260
DB 1359 AATGCACTACCAAGTATTAATCACTACTCAATGCAATTCCTCAAGTAAACACACGGAATCGA 1418
QY 1261 CAAAAAATATGAGAGCAATAGTAAAGATGAGTGGCTTACGCTGCTTGGCAATGAG 1320
DB 1419 CAAAAAATATGAGAGCAATAGTAAAGATGAGTGGCTTACGCTGCTTGGCAATGAG 1478
QY 1321 TGGAACTACTATTAACCAATATGATATGCAATTAAGTGGCTTATGATGAGAGGAGGAGGAG 1380
DB 1479 TGGAACTACTATTAACCAATATGATATGCAATTAAGTGGCTTATGATGAGAGGAGGAGGAG 1538
QY 1381 AGAGTTCTCTAATGTGCGAATCTGTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1539 AGAGTTCTCTAATGTGCGAATCTGTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1598
QY 1441 CATGATGAAAAAGCTTGTACTATGAGATGAGCAAGTCTTATCTTGGCTGCTCC 1500

[illegible]

	BASE COUNT	700 a	518 c	410 g	532 t	ORIGIN
Query Match	99.2%	Score 1983;	DB 1;	Length 2160;		
Best Local Similarity	99.5%	Pred. No. 0;				
Matches 1989;	Conservative	0;	Mismatches	10;	Indels	0; Gaps 0;
QY	1	TAAATCTACGACATATAAAATCAACCTCATTCGTGCTGACCTGGGCTTCTGAACGCCCGCTCAA	60			/gene="pbpia"
DB	159	TAAATCTACGACATATAAAATCAACCTCATTCGTGCTGACCTGGGCTTCTGAACGCCCGCTCAA	218			/transl_start=1
QY	61	TGCCCAAGCTAATGATATATCCACAGATTGGTGAAGCAACGTTTCTATCGAAGACCA	120			/transl_table=1
DB	219	TGCCCAAGCTAATGATATATCCACAGATTGGTGAAGCAACGTTTCTATCGAAGACCA	278			/product="penicillin-binding protein 1A"
QY	121	TGCGTTCTTCGACACACAGGGGGATTGATACCATCCGATCCCTGGGAGCTTTCTTGCGCAA	180			/protein_id="AAPI7256.1"
DB	279	TGCGTTCTTCGACACACAGGGGGATTGATACCATCCGATCCCTGGGAGCTTTCTTGCGCAA	338			/db_xref="GI:6563339"
QY	181	TCTGCAAGCAATTCCTCTCCCAAGTGGATCAACTCTACCCCAACAGTTGAATGTTGAC	240			/translation="MKPKPTILRLIKYISFLSLIVIAIYGVGVFFYYSKAPSLSE
DB	339	TCTGCAAGCAATTCCTCTCCCAAGTGGATCAACTCTACCCCAACAGTTGAATGTTGAC	398			SKIVATSSKLYDNKNOLIADLGSERVVNAQANDIPDLKAVIVSIDHNFPHRGID
QY	241	TTACTTTTCAACTTGTGACTTCGACCAACATATTTCTGTAAAGCTCGAGAAAGCTTGGTT	300			TIRLIGAFRLNQLSLOGSSTLPOQIKILTFESTSDTISKQAEAMLAIOLEK
DB	399	TTACTTTTCAACTTGTGACTTCGACCAACATATTTCTGTAAAGCTCGAGAAAGCTTGGTT	458			ATKQELITYINIKYMSNGNYQMOTAAQYNGKDLNLSLPOLLALAGMQAONQDP
QY	301	AGCGATTGAGTAGAACAACCAAGCAAGCAAGAAATCTTGACCTATATATAATA	360			YSHEAODRNLVLSMKNGYISADQYKAVNPTLTDIGSLQSKASNYPAYMDNL
DB	459	AGCGATTGAGTAGAACAACCAAGCAAGCAAGAAATCTTGACCTATATATAATA	518			KEYINQVEETGYMLTTGMDVYNNNOPEAKHIMDIYNDIDEVAYDDDELQVASTIV
QY	361	GGTCTACATGCTATATGGGAACATATGGAATGACAGACAGACTCAAAACTCTATGTTAA	420			DVSGKTIADLGARHSSNVSPGIGNQVAETNRDMSGTKPIITDAPALREYGVDSAT
DB	519	GGTCTACATGCTATATGGGAACATATGGAATGACAGACAGACTCAAAACTCTATGTTAA	578			IIVHDEPNYPTGNTPNVYNDMGYFNGNTLIYALQOASNVAPAEVFNKGRTEPLN
QY	421	AGACCTCATATAATTAAGTTTACCTCACTATGACCTTGGCTGGGTGAATGCCCTCAGGCAC	480			GLGDIPIYSHNSNAISSNTDESDFKATSEKMAAAAPAGVGTGPKPYIHKVVS
DB	579	AGACCTCATATAATTAAGTTTACCTCACTATGACCTTGGCTGGGTGAATGCCCTCAGGCAC	638			DGSKERSNVGTRAMKETTAAMTEMKATVLSYGTGNALVAILPOGKGTGNTYDE
QY	481	AAACCAATATGACCCCTATTTCACATCCAGAGACAGCCCAAGACCGCGCAAACTTGGCTTT	540			EIEHKITSOFAVPADELPACTYTRKYSAAVMTGYSNRLPLTVGNGLTAAKVVYSMTY
DB	639	AAACCAATATGACCCCTATTTCACATCCAGAGACAGCCCAAGACCGCGCAAACTTGGCTTT	698			LSSESNPEDMNIPEELYNRGEFFVRKNARSTWNSPAQOPPTSESSSSSDSTSOSS
QY	541	ATCTGAAATGAATAATCAAGGCTACATCTGTCTGCTGACAGATATGAAAGCATGTCATATC	600			STTSPSTNNSTTNNNTQOSENTPDOONONPOPAOP
DB	699	ATCTGAAATGAATAATCAAGGCTACATCTGTCTGCTGACAGATATGAAAGCATGTCATATC	758			
QY	601	ACCAATCTGATGAGACACAAAGCTCAAAATCGACAGTAAATTAACCTGCTTACATGGA	660			
DB	759	ACCAATCTGATGAGACACAAAGCTCAAAATCGACAGTAAATTAACCTGCTTACATGGA	818			
QY	661	TAAATCCTCAAGAGATCATCAATCAAGTTGAGAAACCAAGCTTAATCTACTAC	720			

Db 819 TAATTACCTCAAGAGATCATCATCAATCAAGTTGAGAGACAGACGGCTATACCTGCTAC 878
Qy 721 AACTGGATGATGATGCTCTACACAAATGTAGACCAGAAAGCTCAAAAACATCTGTGGATAT 780
Db 879 AACTGGATGATGATGCTCTACACAAATGTAGACCAGAAAGCTCAAAAACATCTGTGGATAT 938
Qy 781 TTACAATACAGACGAATACGTTGGCTATTCAGAGATGAATTGCAAGTCGTTTACCAT 840
Db 939 TTACAATACAGACGAATACGTTGGCTATTCAGAGATGAATTGCAAGTCGTTTACCAT 998
Qy 841 TGTATATGTTTAAAGGAAAGTATGCTGAGTATGAGAGACGCTATCATCAAGTTAA 900
Db 999 TGTATATGTTTAAAGGAAAGTATGCTGAGTATGAGAGACGCTATCATCAAGTTAA 1058
Qy 901 TGTTCCTTCGGAATTTACCAAGACGATAGAAACAAACCGCATGGGATCAACTATGAA 960
Db 1059 TGTTCCTTCGGAATTTACCAAGACGATAGAAACAAACCGCATGGGATCAACTATGAA 1118
Qy 961 ACCGATACAGACTATGCTCTGCTTGAGTAGGGTGTCTACGATTCAACTGCTACTAT 1020
Db 1119 ACCGATACAGACTATGCTCTGCTTGAGTAGGGTGTCTACGATTCAACTGCTACTAT 1178
Qy 1021 CGTTACGATGAGCCCTATTAATACCTGGGCAAAATACTGCTTTTAACTGGGATAG 1080
Db 1179 CGTTACGATGAGCCCTATTAATACCTGGGCAAAATACTGCTTTTAACTGGGATAG 1238
Qy 1081 GGGCTACTTTGGCAACATCATCCTTGCTCAATACGCCCTGCACAAATGCGCAACGTCGCCAGC 1140
Db 1239 GGGCTACTTTGGCAACATCATCCTTGCTCAATACGCCCTGCACAAATGCGCAACGTCGCCAGC 1298
Qy 1141 CGTGAACACTTAAACAAAGGTGGACTACCGCGCCCAAGACTTTCTTAATGCTGTAGG 1200
Db 1299 CGTGAACACTTAAACAAAGGTGGACTACCGCGCCCAAGACTTTCTTAATGCTGTAGG 1358
Qy 1201 AATGACATACCCAGATATTTCACTACTCAAAATGCAATTTCACTAAACACAAACGAATCAGA 1260
Db 1359 AATGACATACCCAGATATTTCACTACTCAAAATGCAATTTCACTAAACACAAACGAATCAGA 1418
Qy 1261 CAAAAAATATGAGACAGATAGTGAAGATGAGTCTGCTTACGCTGCTTGGCAAAATG 1320
Db 1419 CAAAAAATATGAGACAGATAGTGAAGATGAGTCTGCTTACGCTTGGCAAAATG 1478
Qy 1321 TGGACCTACTATTAACCAATGTATATCCATTAAGTCTGCTTGTAGTGGAGTGAANA 1380
Db 1479 TGGACCTACTATTAACCAATGTATATCCATTAAGTCTGCTTGTAGTGGAGTGAANA 1538
Qy 1381 AGAGTTCTCTAATGTCGGAACGTCGCAATGAGGAAGAAACGACAGCCTATATGATACCGA 1440
Db 1539 AGAGTTCTCTAATGTCGGAACGTCGCAATGAGGAAGAAACGACAGCCTATATGATACCGA 1598
Qy 1441 CATGATGAAGACGCTTGTACTTATGAACTTGACGAATGCTTATGCTTGGCTGCC 1500
Db 1599 AATGATGAAGACGCTTGTACTTATGAACTTGACGAATGCTTATGCTTGGCTGCC 1658
Qy 1501 TCAGCCTGTAAGACAGAACCTCTAATACAGACGAGAAATTTGAAACACCATCA 1560
Db 1659 TCAGCCTGTAAGACAGAACCTCTAATACAGATGAGAAATTTGAAACACCATCA 1718
Qy 1561 GACCTCTCAATTTGAGACCGTGAATCTTTGCTGGCTATACGCGTAATATTTCAT 1620
Db 1719 GACCTCTCAATTTGAGACCGTGAATCTTTGCTGGCTATACGCGTAATATTTCAT 1778
Qy 1621 GACCTGATGAGACGCTATTTCAACCGTCTGACACACTTGTAGGCAATGACCTTACGGT 1680
Db 1779 GACCTGATGAGACGCTATTTCAACCGTCTGACACACTTGTAGGCAATGACCTTACGGT 1838
Qy 1681 CGCTGCCAAAGTTTACCGCTCTATATGACTACTGCTGTGAGAGACATCCAGAGA 1740
Db 1839 CGCTGCCAAAGTTTACCGCTCTATATGACTACTGCTGTGAGAGACATCCAGAGA 1898
Qy 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAATAATGGGCTCG 1800
Db 1899 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAATAATGGGCTCG 1958

Qy 1801 TTCTACGTGGAACTACCTGCTCCACAAACACCCCATCACTGAAGTTCAAGCTCATC 1860
Db 1959 TTCTACGTGGAACTACCTGCTCCACAAACACCCCATCACTGAAGTTCAAGCTCATC 2018
Qy 1861 ATCAGATAGTTAACTTCACTACAGTCTAGCTCAACCACTCCACACAAATTAATAGTACGC 1920
Db 2019 ATCAGATAGTTAACTTCACTACAGTCTAGCTCAACCACTCCACACAAATTAATAGTACGC 2078
Qy 1921 TACCAATCCTTACCAATTAATACGCAACCAATCAATCAACCCCTGATTCACAAATACAGA 1980
Db 2079 TACCAATCCTTACCAATTAATACGCAACCAATCAATCAACCCCTGATTCACAAATACAGA 2138
Qy 1981 TCCTCAACCGACCAACCA 1999
Db 2139 TCCTCAACCGACCAACCA 2157

RESULT 7
SPONAN7
LOCUS S. pneumoniae (63915) pona gene for penicillin-binding protein 1a. 2157 bp DNA linear BCT 11-MAY-1995
DEFINITION X67872.1 GI:47417
ACCESSION X67872.1 GI:47417
KEYWORDS penicillin-binding protein 1a; pona gene.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.

REFERENCE 1 (bases 1 to 2157)
AUTHORS Hakenbeck, R.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek. Genetik, Ihnestr. 73, 1000 Berlin 33, FRG
REFERENCE 2 (bases 1 to 2157)
AUTHORS Martin, C., Sibold, C. and Hakenbeck, R.
TITLE Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)
MEDLINE 93010977

FEATURES
source location/Qualifiers

gene 1. .2157
/organism="Streptococcus pneumoniae"
/strain="63915"
/db_xref="taxon:1313"
/clone="pckK1-pckK16"
1. .2157
/gene="pona"
1. .>2157
/gene="pona"

CDS
1. .>2157
/product="penicillin-binding protein 1a"
/protein_id="CA48072.1"
/db_xref="GI:47418"

translation="MNKPTILRLIKYLSIFSLVIAIIVLGGVFFYYKAPSLSE
SKIATVTSRYDKNKOLIDIGSEBRVNAQADIPDLVAIYISIEHREFDHDEK
TIRILGAFRLRLQNSLOGSTLPOQLIKLTFSTSDOTISKABEAMALIOLEK
ATKOEILLYTINKYMSNGNTGKTAQNTYGRKLNLSLQALILAGMPQAPNDP
YSHBAODRNLVLSEKNOGYISADYEKAIVPTTDGLQSLKASNPAYMDTV
KEVINOVEERTGYMLTGTMDVYTVNDOEAKHLMIDYNDDEVAYAPDDEQVASTIV
DVSNGKYIAGLGARHOSNVSFGINOVETVRDGSNPKPTDYAPALEYGIYDSTAT
IVHDKPYNPBETNPVYNMDRGYFGNTILOVALQDSNRVPVETLNVGLABRAETPLN
GLGIDYRSIHNSNLSSTTSDDKRYCASSSKMAAAYAARFNGCTYYKPKMYIRKTVFS
DGSEKFSNNGTIRAIKETATYMTDMKTYLSTGTRNATLAMIPOAKGTSTNTTB
EIEHNIKTSQVAPDELFAGYTRKYSMAVWGYSNRLPLVNGULTVAAKYRSMTY
LSEGSNPDNMIPEGLYRNGEYFVKNGARSTWNSPAPQPPSTESSSSSSSTSSSS
STPSTNNTTNNNTTQGSNTTTPDOONONPQAPQ"

BASE COUNT 699 a 521 c 407 g 530 t
ORIGIN

Query Match 99.1%; Score 1981.4; DB 1; Length 2157;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY	1	TAAATCTACGACAAATAAAATCAATCTGCTGACTTGGGTTCTGAACGCCGGCTCA	60
DB	159	TAAATCTACGACAAATAAAATCAATCTGCTGACTTGGGTTCTGAACGCCGGCTCA	218
QY	61	TGCCAAGCTAATGATATTCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACA	120
DB	219	TGCCAAGCTAATGATATTCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACA	278
QY	121	TGCGTCTTCGACACAGGGGGATTGATACCATCCGTTCTCGGAGCTTCTTCCGAA	180
DB	279	TGCGTCTTCGACACAGGGGGATTGATACCATCCGTTCTCGGAGCTTCTTCCGAA	338
QY	181	TCTGCAAGCAATTCCTCCAGGTGATCACTCACTCAACCAAGTTGATTAAAGTTAC	240
DB	339	TCTGCAAGCAATTCCTCCAGGTGATCACTCACTCAACCAAGTTGATTAAAGTTAC	398
QY	241	TTACTTTTCAACTTCGACTCCGACAGATTTTCTGTAAGGCTCAGGAAGCTTGGTT	300
DB	399	TTACTTTTCAACTTCGACTCCGACAGATTTTCTGTAAGGCTCAGGAAGCTTGGTT	458
QY	301	AGGATTCAGTTAGAACAAAGAACCAACCAAGCAAAATCTTGACTCTATATAATTA	360
DB	459	AGGATTCAGTTAGAACAAAGAACCAACCAAGCAAAATCTTGACTCTATATAATTA	518
QY	361	GGTGTACATGCTAATGGAAGTATGAAATGACAGACAGCTCAAACTCTATGTAA	420
DB	519	GGTGTACATGCTAATGGAAGTATGAAATGACAGACAGCTCAAACTCTATGTAA	578
QY	421	AGACCTCAATATTTAAGTTTACCTCACTAGCTTGGCTGGAGATGCTCAGGACCC	480
DB	579	AGACCTCAATATTTAAGTTTACCTCACTAGCTTGGCTGGAGATGCTCAGGACCC	638
QY	481	AAACCAATATGACCCCTATTCACATCCAGAGACGCCCAACGCCCAAACTTGGTCT	540
DB	639	AAACCAATATGACCCCTATTCACATCCAGAGACGCCCAACGCCCAAACTTGGTCT	698
QY	541	ATCTGAATGAAATATCAAGGCTACATCTCTGTAACAGATGAAAGAGCTCAATAC	600
DB	699	ATCTGAATGAAATATCAAGGCTACATCTCTGTAACAGATGAAAGAGCTCAATAC	758
QY	601	ACCAATTAAGTATGACTACAAAGTCTCAATCAGCAAGTAATACCTGCTTACATGA	660
DB	759	ACCAATTAAGTATGACTACAAAGTCTCAATCAGCAAGTAATACCTGCTTACATGA	818
QY	661	TAACTACTCAAGAGATCATCAAGTTGAGAGAGAAACAGGCTTAACCTACTAC	720
DB	819	TAACTACTCAAGAGATCATCAAGTTGAGAGAGAAACAGGCTTAACCTACTAC	878
QY	721	AACGTGATGATGCTCTACAAATGTAGACCAAGAGCTCAAAAACCTGTGGGATAT	780
DB	879	AACGTGATGATGCTCTACAAATGTAGACCAAGAGCTCAAAAACCTGTGGGATAT	938
QY	781	TTACAATACAGAGCAATAGTTGCTATCCAGAGATGAATTTGCAAGTCCGTTACCAT	840
DB	939	TTACAATACAGAGCAATAGTTGCTATCCAGAGATGAATTTGCAAGTCCGTTACCAT	998
QY	841	TGTTGATGTTTCTAAGGTTAAAGTATGCGCCAGCTAGAGACGCCATCAAGTAA	900
DB	999	TGTTGATGTTTCTAAGGTTAAAGTATGCGCCAGCTAGAGACGCCATCAAGTAA	1058
QY	901	TGTTTCTTGGAAATTAACCAAGAGTAGAAGAAACCAACCGGAGCTGGGATCAATGAA	960
DB	1059	TGTTTCTTGGAAATTAACCAAGAGTAGAAGAAACCAACCGGAGCTGGGATCAATGAA	1118
QY	961	ACCGATACAGACTATGCTGCTGCTGGAGTACGGTGTCTACGATTCACACTGCTACTAT	1020
DB	1119	ACCGATACAGACTATGCTGCTGCTGGAGTACGGTGTCTACGATTCACACTGCTACTAT	1178

QY	1021	CGTTCAGATGAGCCCTTAATACCTGGAGCAATACTCCTGTTTAACTGGGATAG	1080
DB	1179	CGTTCAGATGAGCCCTTAATACCTGGAGCAATACTCCTGTTTAACTGGGATAG	1238
QY	1081	GGGCTACTTTGGCAACATCACCTTGCATAGCCCTGCAACATGCGCAACGTCACAC	1140
DB	1239	GGGCTACTTTGGCAACATCACCTTGCATAGCCCTGCAACATGCGCAACGTCACAC	1298
QY	1141	CGTGAACCTCTAACAAGTGGAGCTCAACCGCGCCCAAGACTTCTTAATGTCATAG	1200
DB	1299	CGTGAACCTCTAACAAGTGGAGCTCAACCGCGCCCAAGACTTCTTAATGTCATAG	1358
QY	1201	AATGCACTAACCAATATCTACTCAATTCCTCAAGTTCCTCAAGTACCAACCGATGCA	1260
DB	1359	AATGCACTAACCAATATCTACTCAATTCCTCAAGTTCCTCAAGTACCAACCGATGCA	1418
QY	1261	CAAAAATATGAGCAATAGTAAGAAATGAGTCTGCTTACGCTGCTTTCGCAATAG	1320
DB	1419	CAAAAATATGAGCAATAGTAAGAAATGAGTCTGCTTACGCTGCTTTCGCAATAG	1478
QY	1321	TGGAATCTACTATAACCAATGTATATCCATTAAGTGTCTTTAGTGAATGGATGAAA	1380
DB	1479	TGGAATCTACTATAACCAATGTATATCCATTAAGTGTCTTTAGTGAATGGATGAAA	1538
QY	1381	AGAGTCTCTAATGTGGAACCTCGTCCATGAGGAACGACGCTATATGATACCGA	1440
DB	1539	AGAGTCTCTAATGTGGAACCTCGTCCATGAGGAACGACGCTATATGATACCGA	1598
QY	1441	CATGATGAAAACAGCTCTGACTTAATGAAATGAGCAAGTGCCTTTCGCTGGCTCC	1500
DB	1599	CATGATGAAAACAGCTCTGACTTAATGAAATGAGCAAGTGCCTTTCGCTGGCTCC	1658
QY	1501	TCAGGCTGTAAACAGCAACCTCTAATATACAGAGCAAGAAATGAAAACCATGAA	1560
DB	1659	TCAGGCTGTAAACAGCAACCTCTAATATACAGAGCAAGAAATGAAAACCATGAA	1718
QY	1561	GACCTCTCAATTTAGAACCTGATGAACTATTCCTGGCTATACGCTAAATATTCAT	1620
DB	1719	GACCTCTCAATTTAGAACCTGATGAACTATTCCTGGCTATACGCTAAATATTCAT	1778
QY	1621	GGCTGTATGAGACAGGCTATTTAAACCGTGTGACACACCTTGTAGGCAATGGCTTACG	1680
DB	1779	GGCTGTATGAGACAGGCTATTTAAACCGTGTGACACACCTTGTAGGCAATGGCTTACG	1838
QY	1681	CGTGGCAAAAGTTTACCGCTCTATGATGACCTTACCTGCTGTAAGGAGCAATCCAGAGA	1740
DB	1839	CGTGGCAAAAGTTTACCGCTCTATGATGACCTTACCTGCTGTAAGGAGCAATCCAGAGA	1898
QY	1741	TTGGAATATACCAAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAATGTCGTCG	1800
DB	1899	TTGGAATATACCAAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAATGTCGTCG	1958
QY	1801	TTTCAAGTGAAGTACCTGCTCCACACACACCCCATCACTGAAGTTCAAGCTCATC	1860
DB	1959	TTTCAAGTGAAGTACCTGCTCCACACACACCCCATCACTGAAGTTCAAGCTCATC	2018
QY	1861	ATCAGATAGTTCACCTTACAGTGTAGCTCAACCACTCAAGCACAAATATAGTAGAC	1920
DB	2019	ATCAGATAGTTCACCTTACAGTGTAGCTCAACCACTCAAGCACAAATATAGTAGAC	2078
QY	1921	TACCAATCTTAACAATATATAGCAACATCAATATCAACCCCTGATCAACAAATACAGAA	1980
DB	2079	TACCAATCTTAACAATATATAGCAACATCAATATCAACCCCTGATCAACAAATACAGAA	2138
QY	1981	TCCTCAACCCAGCAACCA 1999	
DB	2139	TCCTCAACCCAGCAACCA 2157	

RESULT 8
AX11315 2160 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 2048 from Patent WO0123604.

ACCESSION AX111315 GI:13927607
VERSION AX111315.1
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Bergeron,M.G., Boltslot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2048 05-Apr-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
location/Qualifiers
FEATURES
source 1. 2160
/organism="Streptococcus pneumoniae"
/strain="R6"
/db_xref="taxon:1313"
BASE COUNT 697 a 522 c 413 g 528 t
ORIGIN
Query Match 99.1%; Score 1981.4; DB 6; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 198; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 TAAATCTAGACAAATAAATCAATCATCTTGCTGACTGGGCTCTGAAGCGCGCTCAA 60
DB 159 TAAATCTAGACAAATAAATCAATCATCTTGCTGACTGGGCTCTGAAGCGCGCTCAA 218
QY 61 TGCCCAAGCTAATGATATTCACAGATTGGTTAAGCAATCGTTCTATCGAAGCA 120
DB 219 TGCCCAAGCTAATGATATTCACAGATTGGTTAAGCAATCGTTCTATCGAAGCA 278
QY 121 TCGCTTTTGACACAGAGGGGATTTACATCCGTAATCTGGAGCTTTCTGGCGAA 180
DB 279 TCGCTTTTGACACAGAGGGGATTTACATCCGTAATCTGGAGCTTTCTGGCGAA 338
QY 181 TCTGCAAGCAATTCCTCCAGGTGATCACTCAACCACAGTTGATTAGTTGAC 240
DB 339 TCTGCAAGCAATTCCTCCAGGTGATCACTCAACCACAGTTGATTAGTTGAC 398
QY 241 TTACTTTCAACTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGAAAGCTTGGTT 300
DB 399 TTACTTTCAACTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGAAAGCTTGGTT 458
QY 301 AGCGATTCAGTTAGAACAAAAAGCAACCAAGAAATCTTGACCTATATATAATAA 360
DB 459 AGCGATTCAGTTAGAACAAAAAGCAACCAAGAAATCTTGACCTATATATAATAA 518
QY 361 GGCTACATGCTAATGGAAGCAATGGAATGACAGACAGCTCAAAAATCTATGTTAA 420
DB 519 GGCTACATGCTAATGGAAGCAATGGAATGACAGACAGCTCAAAAATCTATGTTAA 578
QY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGAAATGCCTCAGGACC 480
DB 579 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGAAATGCCTCAGGACC 638
QY 481 AAACCAATATGACCCCTATTTACATCCAGAAAGCAAGCCGCGAAGCTTGCTTT 540
DB 639 AAACCAATATGACCCCTATTTACATCCAGAAAGCAAGCCGCGAAGCTTGCTTT 698
QY 541 ATTCGAATGAAAATGAAGGCTACATCTCTGTAACAGATGAGAAAGAGCTCAATAC 600
DB 699 ATTCGAATGAAAATGAAGGCTACATCTCTGTAACAGATGAGAAAGAGCTCAATAC 758
QY 601 ACCAATTAAGTAGACTACAAAGCTCAAAATGACAGATTAATTAATCCCTGCTCATGGA 660
DB 759 ACCAATTAAGTAGAGGCTACAAAGCTCAAAATGACAGATTAATTAATCCCTGCTCATGGA 818
QY 661 TAATTAAGTAGAGGCTACATCAATCAAGTTGAAGAAAGAAAGAGCTTAATCACTATCAC 720
DB 720 TAATTAAGTAGAGGCTACATCAATCAAGTTGAAGAAAGAAAGAGCTTAATCACTATCAC 780

DB 819 TAATTAAGTAGAGGCTACATCAATCAAGTTGAAGAAAGAAAGAGCTTAATCACTATCAC 878
QY 721 AACTGGATGGATGCTCAACAAATGTAGACCAAGAGCTCAAAAGATCTGTGGGATAT 780
DB 879 AACTGGATGGATGCTCAACAAATGTAGACCAAGAGCTCAAAAGATCTGTGGGATAT 938
QY 781 TTACAAATACAGCAATATGCTTGGCTATCCAGAGATGAATGCAAGTCCCTTACCAT 840
DB 939 TTACAAATACAGCAATATGCTTGGCTATCCAGAGATGAATGCAAGTCCCTTACCAT 998
QY 841 TGTGATGTTTCTAAGGTAAGTCAATTCCTGACCTAGGACAGCCATCGTCAAGTAA 900
DB 999 TGTGATGTTTCTAAGGTAAGTCAATTCCTGACCTAGGACAGCCATCGTCAAGTAA 1058
QY 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAACAACCGCAGCTGGGGATCACTATGAA 960
DB 1059 TGTTCCTTGGGAATTAACCAAGCAGTAGAACAACCGCAGCTGGGGATCACTATGAA 1118
QY 961 ACCGATCAGACATATGCTCTGCTTGGAGTACGGGTCTAGCATTCACCTGCTACTAT 1020
DB 1119 ACCGATCAGACATATGCTCTGCTTGGAGTACGGGTCTAGCATTCACCTGCTACTAT 1178
QY 1021 CGTTCAGATGAGCCCTATATACCTACCGGGGACAATATCTGTTTATTAATCGGATAG 1080
DB 1179 CGTTCAGATGAGCCCTATATACCTACCGGGGACAATATCTGTTTATTAATCGGATAG 1238
QY 1081 GGGCTACTTTGGCAACATCACTTGAATACGCCCTGCAACATACGGGAAACGTCGCCAGC 1140
DB 1239 GGGCTACTTTGGCAACATCACTTGAATACGCCCTGCAACATACGGGAAACGTCGCCAGC 1298
QY 1141 CGTGAAGATCTTAACAAGAGTGGAGTCAACCGCGCAAGCTTTCTTAATGGTCTAGG 1200
DB 1299 CGTGAAGATCTTAACAAGAGTGGAGTCAACCGCGCAAGCTTTCTTAATGGTCTAGG 1358
QY 1201 AATGATCAACCAATATTCCTACTCAATATGCCATTTCAAGTAACCAACGATGAGA 1260
DB 1359 AATGATCAACCAATATTCCTACTCAATATGCCATTTCAAGTAACCAACGATGAGA 1418
QY 1261 CAAAAAATATGAGCAAGTAGTAAGAAAGATGGCTGCTTACGCTTGCTTGGCAATGG 1320
DB 1419 CAAAAAATATGAGCAAGTAGTAAGAAAGATGGCTGCTTACGCTTGCTTGGCAATGG 1478
QY 1321 TGAATCTACTATTAACCAATATTCCTACTCAATATGCCATTTCAAGTAACCAACGATGAGA 1380
DB 1479 TGAATCTACTATTAACCAATATTCCTACTCAATATGCCATTTCAAGTAACCAACGATGAGA 1538
QY 1381 AGATTCCTAATGCGGAATGCGGCAATGAGGAAACGACAGCTTATGATGACGA 1440
DB 1539 AGATTCCTAATGCGGAATGCGGCAATGAGGAAACGACAGCTTATGATGACGA 1598
QY 1441 CATGATGAAAACAGTCTTGAATATGGAACCTGACGAAGAAATCTTGGCTGGCTGCC 1500
DB 1599 CATGATGAAAACAGTCTTGAATATGGAACCTGACGAAGAAATCTTGGCTGGCTGCC 1658
QY 1501 TCAGGCTGGTAAACAGAGACCTTAATATACAGAGAGGAAATGAAACACATCAAA 1560
DB 1659 TCAGGCTGGTAAACAGAGACCTTAATATACAGAGAGGAAATGAAACACATCAAA 1718
QY 1561 GACCTCTCAATTTGAGACCTGATGAACCTATTTGGCTGGCTATAGCGGTAAATATTCAT 1620
DB 1719 GACCTCTCAATTTGAGACCTGATGAACCTATTTGGCTGGCTATAGCGGTAAATATTCAT 1778
QY 1621 GCGTGTATGACAGCTATTTCAACCGCTGACACCACTTTGAGCAATGGCCTTACGGT 1680
DB 1779 GCGTGTATGACAGCTATTTCAACCGCTGACACCACTTTGAGCAATGGCCTTACGGT 1838
QY 1681 CGCTGCCAAAGTTAACCGCTCTATGATGACCTACCTGCTGAAGAGCAATCCAGAAAG 1740
DB 1839 CGCTGCCAAAGTTAACCGCTCTATGATGACCTACCTGCTGAAGAGCAATCCAGAAAG 1898
QY 1741 TTGGAATATACCAAGGGGCTCTACAGAAATGAGGAATTCGTTATTTAAATGGTGGCTCG 1800
DB 1899 TTGGAATATACCAAGGGGCTCTACAGAAATGAGGAATTCGTTATTTAAATGGTGGCTCG 1958

OY	1801	TTCACGAGGACATCACTCGTCCACACAACCOCCTCATCACTGAAGTTCAAGTCAATC	1860
Db	1959	TTCTACGTGGAGCTCACTCGTCCACAACCCCCCATCAACTGAAGTTCAAGTCAATC	2018
OY	1861	ATCACATAGTTCAACTTCACAGTGTAGCTCAACCACTCCAAGCACAAATAATAGTAGCANC	1920
Db	2019	ATTCAGATAGTTCAACTTCACAGTGTAGCTCAACCACTCCAAGCACAAATAATAGTAGCANC	2078
OY	1921	TACCAATCCTTAACAATAATAGCGAACAATCAAATTCAAACCCCTGATCAACAATAATCAGAA	1980
Db	2079	TACCAATCCTTAACAATAATAGCGAACAACAATCAAACCCCTGATCAACAATAATCAGAA	2138
OY	1981	TCCCTACACGACGACACCA	1999
Db	2139	TCTTCAACGACGACACCA	2157
RESULT	9		
STRPONA			
LOCUS		3378 bp DNA linear BCT-26-APR-1993	
DEFINITION	STRPONA	Streptococcus pneumoniae penicillin-binding protein (pona) gene,	
	complete cds.		
ACCESSION	M90527		
VERSION	M90527.1 GI:153766		
KEYWORDS	penicillin-binding protein; pona gene.		
SOURCE	Streptococcus pneumoniae (strain R6) DNA.		
ORGANISM	Streptococcus pneumoniae Bacteria; Filimicrobia; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		
REFERENCE	1 (bases 1 to 3378) Martin,C., Briese,T. and Hakenbeck,R. Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae and Streptococcus oralis with high homology to Escherichia coli penicillin-binding proteins 1A and 1B J. Bacteriol. 174, 4517-4523 (1992)		
JOURNAL			
AUTHORS			
TITLE			
MEDLINE			
FEATURES	92325042		
source	Location/Qualifiers		
	1..3378		
	/organism="Streptococcus pneumoniae"		
	/strain="R6"		
gene	/db_xref="taxon:1313"		
	315..3105		
	/gene="pona"		
	315..320		
TATA_signal	/gene="pona"		
CDS	353..949		
	/gene="pona"		
	/function="unknown"		
	/note="orf"		
	/codon_start=1		
	/transl_table=1		
	/protein_id="AAA26955.1"		
	/db_xref="GI:153767"		
	/translation="MNVYPHKVSSOKROTSLSDPKNEANRGMSPFKMINATNDYYLSQ GLATIHKKPEIQLIVOVDPQRSRAKLYEAFROASTDYGVAINGAYIDEVEVETKQ KRALPMKNREYHOHIOHMBOVLAOOGICFVLHFESOOEYLLPAFDLIRFYHQDGKGK SMPLERYEVHELKAGAFPQIPILNYIKELHLGKTR"		
	946..3105		
CDS	/gene="pona"		
	/codon_start=1		
	/transl_table=1		
	/product="penicillin-binding protein"		
	/protein_id="AAA26956.1"		
	/db_xref="GI:153768"		
	/translation="MNKPYLRLIKYISIFSLVIAIYLGGVFFYYYSKADLSSE SKLATTSKIYDNKNLIADIGSERRYNAQANDPIDLVAVISIEDHREFDHGRGIDIE TRLIGAFLNLQSNLSLOGSGALTQQOLIKLYFESTSDOTISRKADEAWLAIQLEQR ATKEILITYINKYVNSNGNGMTAONANYGRDLNNLSPOLALLAGMPAPNOYPD YSHEAADNRNLIYSEMKNQGYTSARDYEKANVPITDGIQSLSASNPAVMNDYILY KEYVNOBEEREGYNLLITGDVYTNTVDQEAOKHLMDITYNDEYVAYPDDELQVASTIV IDVHGKATYAQLGARHOSNVSFGINVAEINRWGSGTKPTTDYPALEGGVYSTATN IVHDENPYNPQTNPVYNWDRVGENTTLQYALAQOSRVPAVENTLNKVGINRATTFPN		

Query Match	Best Local Similarity	Score	DB	Length
Matches 1988; Conservative	99.4%; Pred. No. 0;	0; Mismatches	11; Indels	0; Gaps
0Y	1 TAAATCTACGCAATATAAATCAACTGATTCGTGACTTGGGTTCTGACGCCGCTCA	60		
Db	1104 TAAATCTACGCAATATAAATCAACTGATTCGTGACTTGGGTTCTGACGCCGCTCA	1163		
0Y	61 TGGCCCACTAATGATATTCACAGATTTGGTTAAGGCATTCGTTTATCGAAGACCA	120		
Db	1184 TGGCCCACTAATGATATTCACAGATTTGGTTAAGGCATTCGTTTATCGAAGACCA	1223		
0Y	121 TCGCTTCTTCGACACAGGGGGATGATACCATCCGTTCCGCGGAGCTTCTTCGCGCA	180		
Db	1224 TCGCTTCTTCGACACAGGGGGATGATACCATCCGTTCCGCGGAGCTTCTTCGCGCA	1283		
0Y	181 TCTGCAAGCAATTCCTCCAGGTTGATCACTCCACCCACAGTTGATTAAAGTTGAC	240		
Db	1284 TCTGCAAGCAATTCCTCCAGGTTGATCACTCCACCCACAGTTGATTAAAGTTGAC	1343		
0Y	241 TTACTTTTCACTTCGACTTCGCCAGCACAATCTTCTGTAAAGCTCGAAGAGCTGGTT	300		
Db	1344 TTACTTTTCACTTCGACTTCGCCAGCACAATCTTCTGTAAAGCTCGAAGAGCTGGTT	1403		
0Y	301 AGCGATTCAGTTAGCAACAAAAAGCAACCAACGAATTCCTGACCTCTATTAATAA	360		
Db	1404 AGCGATTCAGTTAGCAACAAAAAGCAACCAACGAATTCCTGACCTCTATTAATAA	1463		
0Y	361 GGTCTACATGCTTAATGAGGAACATATGAAATCGACAGCAGCTCAAACTACTATGTAA	420		
Db	1464 GGTCTACATGCTTAATGAGGAACATATGAAATCGACAGCAGCTCAAACTACTATGTAA	1523		
0Y	421 AGACCTCAATTAATTAATTTACCTCACTTGAACCTTGGCTGGAAAGCCTCAGGCACC	480		
Db	1524 AGACCTCAATTAATTAATTTACCTCACTTGAACCTTGGCTGGAAAGCCTCAGGCACC	1583		
0Y	481 AAACCAATATGACCCCTATTCACATCCAGAGACAGCCCAACCGCGCAACTTGGTCTT	540		
Db	1584 AAACCAATATGACCCCTATTCACATCCAGAGACAGCCCAACCGCGCGCAACTTGGTCTT	1643		
0Y	541 ATCTGAAATGAAAAATVCAAGGCTACATCTCTGCTGAACAGTATGAAAAAGCAGTCAATAC	600		
Db	1644 ATCTGAAATGAAAAATVCAAGGCTACATCTCTGCTGAACAGTATGAAAAAGCAGTCAATAC	1703		
0Y	601 ACCAATTTCTGATGAGCTACACAAAGTCTCAATACGAAGTAAATTAACCTGCTTACATGGA	660		
Db	1704 ACCAATTTCTGATGAGCTACACAAAGTCTCAATACGAAGTAAATTAACCTGCTTACATGGA	1763		
0Y	661 TAAATTAACCTCAAGGAAGTCATCAATCAAGTTGAAGAAGAAACAGGCTTAATCACTAC	720		
Db	1764 TAAATTAACCTCAAGGAAGTCATCAATCAAGTTGAAGAAGAAACAGGCTTAATCACTAC	1823		
0Y	721 AACTGGGATGATGCTCTACACAAATGTAGACCAAGACCTCAAAAACATCTGTGGGATAT	780		
Db	1824 AACTGGGATGATGCTCTACACAAATGTAGACCAAGACCTCAAAAACATCTGTGGGATAT	1883		
0Y	781 TTACAAATACAGACGAATACGTTGCTATTCACAGAGATGAATTTGCAAGTGGCTTACCAT	840		
Db	1884 TTACAAATACAGACGAATACGTTGCTATTCACAGAGATGAATTTGCAAGTGGCTTACCAT	1943		
0Y	841 TGTGATATGTTTCTAAGGTTAAAGTCATGGCCAGCTAGAGACAGCCATCAGTCAAGTAA	900		
Db	1944 TGTGATATGTTTCTAAGGTTAAAGTCATGGCCAGCTAGAGACAGCCATCAGTCAAGTAA	2003		

QY	901	TGTTTCCCTCGGAAATTAAACCAAGAGTAGAAACAAACCGGAGCTGGGGATCAACTATGGA	960
Db	2004	TGTTTCCCTCGGAAATTAAACCAAGAGTAGAAACAAACCGGAGCTGGGGATCAACTATGGA	206
QY	961	ACCGATCAAGACATGTGCTCTCGCTTGGAGTAGTACGGTGTCTACGATTCAACTGCTACTAT	1021
Db	2064	ACCGATCAAGACATGTGCTCTCGCTTGGAGTAGTACGGTGTCTACGATTCAACTGCTACTAT	212
QY	1021	CGTTTCAGAGTAGAGCCCTTAATTAACCTACCTGGGACAAATACTCTGTTTATTAACCTGGATAG	108
Db	2124	CGTTTCAGAGTAGAGCCCTTAATTAACCTACCTGGGACAAATACTCTGTTTATTAACCTGGATAG	218
QY	1081	GGGGCTACTTTGGCAACATCAACCTTGGCAATTAAGCCCTTSCAACATTCGGGAAAGCTCCACAG	1144
Db	2184	GGGGCTACTTTGGCAACATCAACCTTGGCAATTAAGCCCTTSCAACATTCGGGAAAGCTCCACAG	224
QY	1141	CGTGGAAACTCTAAACAAGGTGCGAGCTCAACCGCGCCACAGACTTCTCTAAATGTCCTAGG	1201
Db	2244	CGTGGAAACTCTAAACAAGGTGCGAGCTCAACCGCGCCACAGACTTCTCTAAATGTCCTAGG	230
QY	1201	AATGCACTACCCANGTATTCACTAATCTCAATGCCATTTCAGATTAACACACGGAATCAGA	126
Db	2304	AATGCACTACCCANGTATTCACTAATCTCAATGCCATTTCAGATTAACACACGGAATCAGA	236
QY	1261	CAAAAAATATGAGACCAAGTATGTAAGAAAGATGGGTCGCTTACGTCGCTTTGSCAAATGG	132
Db	2364	CAAAAAATATGAGACCAAGTATGTAAGAAAGATGGGTCGCTTACGTCGCTTTGSCAAATGG	242
QY	1321	TGGAACCTACTATTAACCAATGTATATCATTAAGTCGTCCTTTAGTAGTAGGAGTAGTAAAA	1380
Db	2424	TGGAACCTACTATTAACCAATGTATATCATTAAGTCGTCCTTTAGTAGTAGGAGTAGTAAAA	248
QY	1381	AGAAGTCTCTAATGTGGAACCTGCTGCCATGAAGAAAGACAGAGCTATATGTATGACCGA	1444
Db	2484	AGAAGTCTCTAATGTGGAACCTGCTGCCATGAAGAAAGACAGAGCTATATGTATGACCGA	254
QY	1441	CATGATGAAAAACATCTTGACTTATGGAACATGGACGGAAGCCCTATCTGCTTGGCTGGCC	1500
Db	2544	CATGATGAAAAACATCTTGACTTATGGAACATGGACGGAAGCCCTATCTGCTTGGCTGGCC	260
QY	1501	TCAGAGCTGTAATAAACAGAGAACCTTACTATACAGAGAGGAATTTGAAACACACATCAA	1561
Db	2604	TCAGAGCTGTAATAAACAGAGAACCTTACTATACAGAGAGGAATTTGAAACACACATCAA	266
QY	1561	GACCTCTCAATTTGTAGCACCTGATGAACATAATTGCTGGCTATAGCGGTAAATATTCAT	1621
Db	2664	GACCTCTCAATTTGTAGCACCTGATGAACATAATTGCTGGCTATAGCGGTAAATATTCAT	272
QY	1621	GGCGTATGAGCAGGCTATTTCTAAACGTCGTGACACCACTGTGAGGAATGGGCTTACGGT	1680
Db	2724	GGCGTATGAGCAGGCTATTTCTAAACGTCGTGACACCACTGTGAGGAATGGGCTTACGGT	278
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGATGAGACCTACTGTCGTAAGAGAGCAATTCAGAA	1744
Db	2784	CGCTGCCAAAGTTTACCGCTCTATGATGAGACCTACTGTCGTAAGAGAGCAATTCAGAA	284
QY	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGTCCTCG	1800
Db	2844	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGTCCTCG	290
QY	1801	TTTCTAGCTGGAACCTCACTCTCTCCACAAACCCCATCAACGAAAGTTTCAAAGCTCAATC	1861
Db	2904	TTTCTAGCTGGAACCTCACTCTCTCCACAAACCCCATCAACGAAAGTTTCAAAGCTCAATC	296
QY	1861	ATCAGATAGTTCAACTTCACAGTCTAGCTTAACCACTCCAGACCAAAATATATAGTACGAC	1920
Db	2964	ATCAGATAGTTCAACTTCACAGTCTAGCTTAACCACTCCAGACCAAAATATATAGTACGAC	302
QY	1921	TACCAATCTTAACAATTAATACGCAACATCAATATACCAACCCCTGATCAACAATAATCAGAA	1980
Db	3024	TACCAATCTTAACAATTAATACGCAACATCAATATACCAACCCCTGATCAACAATAATCAGAA	308
QY	1981	TTCTTCAACGACACACCA 1999	

Db	3084	TCTCTACGACGACACCA	3102
RESULT 10	AE008414	10313 bp	DNA linear BCT 13-SEP-2001
LOCUS	AE008414/c		
DEFINITION	Streptococcus pneumoniae R6 section 30 of 184 of the complete genome.		
ACCESSION	AE008414	AE007317	
VERSION	AE008414.1	GI:15457885	
KEYWORDS			
SOURCE			
ORGANISM	Streptococcus pneumoniae R6. Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		
REFERENCE	1 (bases 1 to 10313) Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., Dehoff,B.S., Estrem,S., Filtz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmour,R., Glass,J.S., Khoja,H., Kraft,A., Lagace,R., Leblanc,D.J., Lee,L.N., Lefkowitz,E.J., Lu,J., Matsushima,P., McAhren,S., McInney,M., Mcleaster,K., Mundy,C., Nicas,T.I., Norris,F.H., O'Garra,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.R., Rostock,P.R. Jr., Skatrud,P.L. and Glass,J.I.		
REFERENCE	2 (bases 1 to 10313) Genome of the Bacterium Streptococcus pneumoniae Strain R6 J. Bacteriol. 183 (19), 5709-5717 (2001)		
REFERENCE	11544234		
REFERENCE	2 (bases 1 to 10313) Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., Dehoff,B.S., Estrem,S., Filtz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmour,R., Glass,J.S., Hann,A., Khoja,H., Kraft,A., Lagace,R., Leblanc,D.J., Lee,L.N., Lefkowitz,E.J., Lu,J., Matsushima,P., McAhren,S., McInney,M., Mcleaster,K., Mundy,C., Nicas,T.I., Norris,F.H., O'Garra,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.R., Rostock,P.R. Jr., Skatrud,P.L. and Glass,J.I.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-JUL-2001) Infectious Diseases Research, Eli Lilly and Company, Lilly Research Labs, Indianapolis, IN 46285-0438, USA		
FEATURES	Location/Qualifiers		
SOURCE	1. 10313 /organism="Streptococcus pneumoniae R6" /strain="R6" /db_xref="taxon:171101" complement(114..2273) /gene="pbpa" /note="sp0329" complement(114..2273) /gene="pbpa" /EC_number="2.3.2.-" /EC_number="2.4.1.129" /codon_start=1 /transl_table=11 /label=sp0329 /product="Penicillin-binding protein 1A" /protein_id="AK99133.1" /db_xref="GI:15457886" /translation="MNKPTLLIKLKLISFLVIAIYLGGVFFYYSKAPLSLESKVATTSKTYDNKNQGLADIGSERVNAQANDIPDLKAVISJEDHFEFHRKIDATIIELAFNLNOSLOGSALTOOLIKLTFSTSDTISRKAQEMALAIQLBOKATIOELITLYINKVYNSNGSGYMGQTAQNYGYGDNLNLSLPOLALLAGMQAONQYDKYSPDAQDRNLVLSMKNOGYISAQVRYKANTPTTOSLSKASNPAYMDVLTKEVINOVEETGYNLTTGADVTVTNDQEKRLMDITNTDEYATVPDDELGYASTIVDVSNGVYIQLGARHSSNVSGFINQAVETNRDWSGTMPTITDYAPALEGYESTATIVHDEYVYNGTNTPYVNMDSRGFGHITLQALQOSRNPVAVETLKNVGLINRAKFTLNGIADIDPSYINSAISNTSEDSKRYGASKEAAVAAFAANGTYKKIPMYIKHVESDGEKPEFNSNGYFRAMKETTYATYMTDMKTYLISYGTGRNATLAWLPQAGKGTSNYDEEINNHKTQSFVAPDELFCYTKRYISMAVWTGYSNRLPLVNGGLTYAAKYVSNMTYLSGSPEDMNIPGGLYRNGEYFVKNGARVTSSVPAQDPPTSSSSSSSDSTSS		
gene			
CDS			

gene
STPSTNNSTTTNNPNNTTQOSNTTTPDOONONPQAPQ"
/complement(2270. .2866)
/gene="recu"
/note="spr0330"
/complement(2270. .2866)
CDS
/gene="recu"
/note="putative"
/codon_start=1
/transl_table=1
/label_spr0330
/product="Recombination protein u"
/protein_id="AAK9134.1"
/db_xref="GI:15457887"
/translation="MVNYPKVSOKROTSLOPKNFANRGMSEKMINATNDYLSO
GLAVIHKKPTPIQIVQYDYPORSRAKIYEAFRQASTDYSGYNGYITDEVEKETKO
KRAIPMKNFHPHQIOMEQVLAQOQICVLLHFSSQOETYLPAFDLIRFYHODGOK
SMPELYIREYGEIKAGAFPOIPYLVNLIKELLGKTR"
2932. .3459
CDS
/gene="spr0331"
2932. .3459
/gene="spr0331"
/codon_start=1
/transl_table=1
/label_spr0331
/product="Conserved hypothetical protein"
/protein_id="AAK9135.1"
/db_xref="GI:15457888"
/translation="MATALVGYSAFDLGSFSDKDPRLKLIKAIKRDLEAMADGVS
WLVTGSLGEYVWLEVAQEMKTEYGLATIFAFETIHEGNNMGNOMLSRFGQVDF
VKVAYPRHEKGOLRDYQOQFLENTSSYLEFDEENETKLAVFYQKKMNQSDYFIKRL
TFQDLNEAFNEFSEN"
3517. .3858
gene
/gene="spr0332"
3517. .3858
/gene="spr0332"
/codon_start=1
/transl_table=1
/label_spr0332
/product="Conserved hypothetical protein"
/protein_id="AAK9136.1"
/db_xref="GI:15457889"
/translation="MERMASIIFSAKDIFEOFGREYRGYKVEVDFLDVYKDYE
TYALVLSLROELADLKEELRKPSPVQAEPLAATTSMTNFIDILKRLNLEKEY
FGKOILDNSD"
3871. .4274
misc_RNA
/gene="rnpb"
/product="Ribonuclease P - RNA component"
3871. .4274
gene
/gene="rnpb"
4329. .5501
CDS
/gene="spr0333"
4329. .5501
/gene="spr0333"
/codon_start=1
/transl_table=1
/label_spr0333
/product="Conserved hypothetical protein"
/protein_id="AAK9137.1"
/db_xref="GI:15457890"
/translation="MODCKKEFNLIATYAGLAEVAVGRVRELGYDCOVENGVRP
OGVRAIITENMLRADRIKIIVGTFPAKTFEELFGCVRLDMENLIPDARFPIK
AKCVKSLHNEPSVQALSKAVAKLQKHARPGVPLMGPEPKLEVSILKQVATY
MIDITGSSLEKRGYRTERKGAPIKENMAAAILQSLNNYPDKPLIDPTGSGTFCIEAV
MARKMAAGLERSFAPEKMWISDLIOEVTEAKKVERLELDIMGCDIDARMEI
AKANAQAGVAGDITFKOMRVODRSRKINVIINSPYGERLSDDAGVTKLVAMGO
VEAPLTKMSKFIILSDAEFESKYSQADKKKKLNLGLKVLVLYFFGQRYKQEVK"
5514. .6908
gene
/gene="spr0334"
5514. .6908
/gene="spr0334"
/codon_start=1
/transl_table=1
/label_spr0334

/product="Conserved hypothetical protein"
/protein_id="AAK9138.1"
/db_xref="GI:15457891"
/translation="MSKRRNRHKKAEQEPDPDEAKELTVGAIRKNEVEAGVLP
EDSILDKYKQKRODEIADKFAIRQYKEEVEEQSLDOLLQEMKEAVEKSEASEEV
PSSBILPLPLDDEEGDLPDLDDENPTMEVEVEEQNLRLDQDESKKSGF
ILYVALVSVIICVAVYVROVARSTKEIETSOSTNODVDQFNLLYAPFTNSN
KTKALNSQFDKLSQKTLDDLEGRREHTLAKSYDSLATQIKAIQDYNADQPEKPAIY
DGLVDITNAKAKSDAKFTDIKTGNTELDKVLKALISLGSQQTSTSSSSQTSSSSS
QASSNTSEPKSSNETRSRSREVMGLSAGVAVQVSASRAVANQSDAIDSSNSAW
DFADGVLEQIILATSKRSRGYITGDQYILERVNIIVNGNGYNNLYKPKGYTLFLLNCKTGY
FVGNAGHAADLDY"
6984. .8408
gene
/gene="gnd"
/note="spr0335"
6984. .8408
CDS
/gene="gnd"
/EC_number="1.1.1.44"
/codon_start=1
/transl_table=1
/label_spr0335
/product="6-phosphogluconate dehydrogenase"
/protein_id="AAK9139.1"
/db_xref="GI:15457892"
/translation="WTKANPEYGVGAWVRMLAINIESRGYTVAILNRSKRETDVIA
CHPEKNFVSTDESFEVNSIEKPRIMLMVQGPETDFTIQALLPHLDKGLIDIGN
TFYKDTIRNRELANSQINFIGTGVSGEGALGSPSLPESGQAEVAVLADVLEIS
AKADEDGKPCVYTIIGPDAGHYKVMHNGIIEYGMQLIASEYDLMQHLGLSABEMAR
IFETMNKELSELVLEITADILSRKDDGODGPIVDYILDAAGNKGTQKWTQSOSLDL
GPILSITLESVFARYISTYKEERYHASKVLPKPAAPNEGGKATLIEKIRQALYESKI
ISYNOGPAOLRVASKENMNLPFADIASIMDGCITIRSRPQKLTDAIINRADIANLL
IDBYFLDYTKAQIQQARDVVALAVQAGPVTFSAITYFISADIPANLIDQORD
YFGAHTYQKDKESTFHYSWYDEK"
8420. .9109
gene
/gene="csr"
8420. .9109
CDS
/gene="csr"
/note="putative"
/codon_start=1
/transl_table=1
/label_spr0336
/product="response regulator"
/protein_id="AAK9140.1"
/db_xref="GI:15457893"
/translation="WGRKILLEKERNLAHFLSELQKEYRVLDVEGOKALSMAIQ
TDVYLILNVNLGDMQODFAEKLSRTKPAVIMILDHMEDLOELEVEYORFAVSYIY
KPVLIENVARISALIFRGKDPIDIOHCSMKKPYRYNRIRIDVEHHTYRGEMIALIR
REYDILATLMSKRVLTREQLLESVWKTESATETNIYDVYIRIRLSKLDVAGOKSYIK
TVRGVGYTMOE"
9207. .10223
gene
/gene="cbpf"
/note="spr0337"
9207. .10223
CDS
/gene="cbpf"
/codon_start=1
/transl_table=1
/label_spr0337
/product="Choline-binding protein F"
/protein_id="AAK9141.1"

Query Match 99.1%; Score 1981.4; DB 1; Length 10313;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 61 TGCCCAAGCTAATGATATTCACAGATTTGGTAAAGCAATCTTCTATGAGACCA 120
2115 TAAATCTACGCAATATAAATCAACTATGCTGCTGAGCGCGGCTCAA 2056
2055 TGCCCAAGCTAATGATATTCACAGATTTGGTAAAGCAATCGTTCTATGAGACCA 1996

OY	121	TCCTCTTGACACACAGGGGGATGTTATACATCCGATCCCGGGAGCTTCTGCGCA	180
OY	121	TCCTCTTGACACACAGGGGGATGTTATACATCCGATCCCGGGAGCTTCTGCGCA	180
Db	1995	TGCGCTCTTGACACACAGGGGGATGTTATACATCCGATCCCGGGAGCTTCTGCGCA	1936
OY	181	TCGCAAGCAATTCCTCCACAGGTGATCACTCCACCAACAGTATTAAAGTGC	240
Db	1935	TCGCAAGCAATTCCTCCACAGGTGATCACTCCACCAACAGTATTAAAGTGC	1876
OY	241	TTACTTTTCAACCTTCGACCTTCGACACGAGCTATTCTCGTAAGGCTCAGGAAGCTTGTT	300
Db	1875	TTACTTTTCAACCTTCGACCTTCGACACGAGCTATTCTCGTAAGGCTCAGGAAGCTTGTT	1818
OY	301	AGCGATTAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTGACCTACTATATAAATTA	360
Db	1815	AGCGATTAGTTAGTAACAAAAAGCAACCAAGCAAGAAATCTGACCTACTATATAAATTA	1756
OY	361	GGTCATCATGTCTATAGGGAATTTGAAATGACAGACAGAGCTCAAAATCTATATGTAA	420
Db	1755	GGTCATCATGTCTATAGGGAATTTGAAATGACAGACAGAGCTCAAAATCTATATGTAA	1696
OY	421	AGACCTCAATATTTTAAAGTTTAACTCTCAGTTAGCCTTGCTGGCTGAATGCTCAGGACCC	480
Db	1695	AGACCTCAATATTTTAAAGTTTAACTCTCAGTTAGCCTTGCTGGCTGAATGCTCAGGACCC	1636
OY	481	AAACCAATATACCCCTTATTCATATCCAGAGAGCCCAAGACCCGCGAAATTTGGTCTT	540
Db	1635	AAACCAATATATCCCTTATTCATATCCAGAGAGCCCAAGACCCGCGAAATTTGGTCTT	1576
OY	541	ATTCGAATTAAGAAATACAAAGGCTCATCTGCTGTAACAGTATGAAAGCAAGTAAATAC	600
Db	1575	ATTCGAATTAAGAAATATCAAGGCTCATCTCTGCTGAACAGTATGAAAGCAAGTAAATAC	1516
OY	601	ACCAATTAAGTATGAGTACAAAGTCTCAAAATCAGCAAGTAAATTAACCTGCTTACATGGA	660
Db	1515	ACCAATTAAGTATGAGTACAAAGTCTCAAAATCAGCAAGTAAATTAACCTGCTTACATGGA	1456
OY	661	TAAATTAAGTATGAGTACAAAGTCTCAAAATCAGTAAAGAAAGCAAGCTATTAACCTACTAC	720
Db	1455	TAAATTAAGTATGAGTACAAAGTCTCAAAATCAGTAAAGAAAGCAAGCTATTAACCTACTAC	1396
OY	721	AACTGGATGATGATCTACACAAATGATAGACAAAGGCTCAAAAACATCTGTGGGATAT	780
Db	1395	AACTGGATGATGATCTACACAAATGATAGACAAAGGCTCAAAAACATCTGTGGGATAT	1336
OY	781	TTACAATACAGAGCAATACGTTGGCTATCCAGACGATGAAATGCAAGTGCCTTACCAT	840
Db	1335	TTACAATACAGAGCAATACGTTGGCTATCCAGACGATGAAATGCAAGTGCCTTACCAT	1276
OY	841	TGTTGATGTTTCTTAACGGTAAAGTCATATGCCACGCTAGAGACCGCATCACTCAAGTAA	900
Db	1275	TGTTGATGTTTCTTAACGGTAAAGTCATATGCCACGCTAGAGACCGCATCACTCAAGTAA	1216
OY	901	TGTTTCCCTTGGGAATTAACCAAGAGTAGAAACAAACCGGACTGGGGTCAACATATGA	960
Db	1215	TGTTTCCCTTGGGAATTAACCAAGAGTAGAAACAAACCGGACTGGGGTCAACATATGA	1156
OY	961	ACCGATCACAGACTATGCTCTGCTGGAGTAGAGTGTCTACGATTCGAATGCTACTAT	1020
Db	1155	ACCGATCACAGACTATGCTCTGCTGGAGTAGAGTGTCTACGATTCGAATGCTACTAT	1096
OY	1021	CGTTACAGATAGCCCTATAATACCTGGGACAAATACTCCTGTTTATTAAGTGGATAG	1080
Db	1095	CGTTACAGATAGCCCTATAATACCTGGGACAAATACTCCTGTTTATTAAGTGGATAG	1036
OY	1081	GGGCTACGTTGGCAACATCACTTTGCAATACGCTCGCAACAAATGGCGAAAGTCCGACG	1140
Db	1035	GGGCTACGTTGGCAACATCACTTTGCAATACGCTCGCAACAAATGGCGAAAGTCCGACG	976
OY	1141	CGTGGAAATCTTAAACAAGGTGCGACTCAACCGCGCAAGACTTTTCTTAAATGCTATAG	1200
Db	975	CGTGGAAATCTTAAACAAGGTGCGAGTCAACCGCGCAAGACTTTTCTTAAATGCTATAG	916
OY	1201	AATGCACTACCCAGTATCTACTACTCAATGCGCAATTTCAAGTAAACACACCGAATTCAGA	1260

Db	Accession	Length	Strain	Species
Db	915	117	ATTCGACTACCCAGATATTCACTACTCAATGCGATTTAAAGTAAACACACCGAATCAGA	856
Qy	1261	1320	CAAAAATATGAGGCAAGTAGTAAAGATGGCTGCTGCTTACGCTTGCCTTTGCCAATGG	1320
Db	855	796	CAAAAATATGAGGCAAGTAGTAAAGATGGCTGCTGCTTACGCTTGCCTTTGCCAATGG	796
Qy	1321	1380	TGGAACTTACTATTAACCAATGATATATCATATTAAGTCGCTTTAGTATGGGAGTGA	1380
Db	795	736	TGGAACTTACTATTAACCAATGATATATCATATTAAGTCGCTTTAGTATGGGAGTGA	736
Qy	1381	1440	AGAGTCTCTATATGTCGGAAGTCTGGCCATGGAAGGAAGCAAGCCTATATGATGACGA	1440
Db	735	676	AGAGTCTCTATATGTCGGAAGTCTGGCCATGGAAGGAAGCAAGCCTATATGATGACGA	676
Qy	1441	1500	CATGATGAAGAAAGTCTTACTTATGGAAGTGGAGAAATGGCTATCTTGGTGGCTCC	1500
Db	675	616	CATGATGAAGAAAGTCTTACTTATGGAAGTGGAGAAATGGCTATCTTGGTGGCTCC	616
Qy	1501	1560	TCAGGCTGTAAACAGAGACTCTTACTATACAGACGAGAAATTTGAACACATCA	1560
Db	615	556	TCAGGCTGTAAACAGAGACTCTTACTATACAGACGAGAAATTTGAACACATCA	556
Qy	1561	1620	GACCTCTCAATTTGATGACACCTGATGAACATATTTGCTGGCTATAGCGTAATTTCAAT	1620
Db	555	496	GACCTCTCAATTTGATGACACCTGATGAACATATTTGCTGGCTATAGCGTAATTTCAAT	496
Qy	1621	1680	GGCTGATGAGACGCTATTTACCGTGTGACACACTTGTAGGCAATGGCCTTACGT	1680
Db	495	436	GGCTGATGAGACGCTATTTACCGTGTGACACACTTGTAGGCAATGGCCTTACGT	436
Qy	1681	1740	CGCTGCCAAGTTTACCGCTGTATGATGACCTTCTGTGAGGAGCAATTCACAGAGA	1740
Db	435	376	CGCTGCCAAGTTTACCGCTGTATGATGACCTTCTGTGAGGAGCAATTCACAGAGA	376
Qy	1741	1800	TTGGAATATACAGAGGGGCTCTACAGAAATGGAGATTCGATTTAAATAATGGGCTCG	1800
Db	375	316	TTGGAATATACAGAGGGGCTCTACAGAAATGGAGATTCGATTTAAATAATGGGCTCG	316
Qy	1801	1860	TTCTACGTGGAAGTACCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTCATC	1860
Db	315	256	TTCTACGTGGAAGTACCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTCATC	256
Qy	1861	1920	ATCAGATAGTCACTTACAGTCTAGTCTCAACCACTCCAGACACAATATATGATGAC	1920
Db	255	196	ATCAGATAGTCACTTACAGTCTAGTCTCAACCACTCCAGACACAATATATGATGAC	196
Qy	1921	1980	TTCGAATCTTACATTAATAGCAACAATCAATCAACCCCTGATTCACAATAATCAGAA	1980
Db	195	136	TTCGAATCTTACATTAATAGCAACAATCAATCAACCCCTGATTCACAATAATCAGAA	136
Qy	1981	1999	TCCCTACACGACACAACCA 1999	
Db	135	117	TCCCTACACGACACAACCA 117	

Db	1659	ACAACAGAGTAAACAGAGTACTTCTAAGTATATCTGACGAGAAATTTGAAATATATATCA	1718
Qy	1561	GACCTCTCAATTTTGAGCACCTGATGAACATAATTTGCTGGCTATPACGCGTAATATTCAT	1620
Db	1719	GAACTAGTGTACGATAGCTCCAGATGAATGTTTGTAGGGTATPACCCGCAATATGCAAT	1778
Qy	1621	GGCTATATGAGACAGGCTATTTCAACCGTGTGACACACTTATGAGCAATGGCGTTACGCT	1680
Db	1779	GGCTATATGAGACAGGCTATTTCAACCGTGTGACACACTTATGAGCGATGGCGTTACGCT	1838
Qy	1681	GCGTCCCAAGTTTACCGCTCTATGATGACCTACCTGTCGTGGAAGGAAATCCAGAGA	1740
Db	1839	CGGCTCTAAAGTTTACCGCTCTATGATGACCTACCTGTCGTGGAAGGAAATCCAGAGA	1898
Qy	1741	TTGGAATATACCAAGGGGCTCTACAGAAATGGAATTCGATTTTAAAAATGCGTCTCG	1800
Db	1899	CTGGAATATACCAAGGGGCTCTACAGAAATGGAATTCGATTTTAAAAATGCGTCTCG	1958
Qy	1801	TTCTACGTGGAACCTCAACCTGTCACCAACACACCCCATATCACTGAAAGTCTCAAGCTCATC	1860
Db	1959	TTCTACGTGGAACCTCAACCTGTCACCAACACACCCCATATCACTGAAAGTCTCAAGCTCATC	2018
Qy	1861	ATTCAGATAGTTAACTTCATGCTAGGCTCAACCACTCCAAACCAAAATATATAGTACGAC	1920
Db	2019	ATTCAGATAGTTAACTTCATGCTAGGCTCAACCACTCCAAACCAAAATATATAGTACGAC	2078
Qy	1921	TACCAATCTTAAACAATATATACGCAACAATCAAAATCAACCCCTGATCAACAATAATCAGAA	1980
Db	2079	TACCAATCTTAAACAATATATACGCAACAATCAAAATCAACCCCTGATCAACAATAATCAGAA	2138
Qy	1981	TGCTCAACGACGACAACCA	1999
Db	2139	TGCTCAACGACGACAACCA	2157

```

CDS
1. 2160
/gene="pbp1a"
/function="peptidoglycan transglycosylase-transpeptidase"
/feature_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="AAPD3067.1"
/db_xref="GI:5410459"
/translation="MNKPTLRLIKYLSISFLSLVIAIVLGGVFFVYKSPALSE
SKLVAATTSKRYDKNKLIDLSESRVSNQANOIPDLVAKAIYSIDHFFPRGID
SIRLIGAFRLRLOSLSLOGSTLQOLIKLTLEFSTSPDITSRKAGEAMFLIQLEQ
ATROEILITYYINRKYMSNGNMGQTAQNAVYVGRNLNLSLQALAAQAPQAPQ
KEVSHPEAQDRNRLVLESKNNGITISAOCEKAVTPTDGLQSLKASNTPAYMDNL
KEVSHNOVEQETGYNLTTGMDVYVNVQDEKAKHLMIDYINSDQVSYPTDDLOVAATV
DVNSGKVIYALQGARHQASNVSEFTGNQAVETNRDQSSMKPITDAPALEGVYDSTAT
MVNDIPYNPCTSPRYVNMDBRAYFGNLTLYALQSSNVAYETLNVGLDRAKTEFL
GLGIDYESMVHYANAISNTTESNKKQYQASSKMAAAFAANGSTYKPYIKHVVES
DGSKEESNPYTRAMKETTAVMATDMKATVLYCTGTGAYLPMPLPGAGKTGSYDE
EVENHIKNTGVADPEMDEVCYTRKYSNAVMYGSNRLPIVGDGEFLAAVYRSMYIY
LSRESNPEDNAPICGLYRNCPEFYVRKNAARWMSDPADQCPSTRESSSSSSDSTSSQ
STTSTSTNNSTTNNNNNTQDSNTTTPDQANONPQAPQ"
BASE COUNT 693 a 494 c 407 g 566 t
ORIGIN

```

```

RESULT 12
AF139884
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene

AF139884 2160 bp DNA linear BCT 26-AUG-1999
Streptococcus pneumoniae strain PO-273 penicillin-binding protein
1a (pbp1a) gene, complete cds.
AF139884
AF139884.1 GI:5410458

Streptococcus pneumoniae.
Streptococcus pneumoniae
Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2160)
Cofey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
Serotype 14 variants of the Spanish penicillin-resistant serotype
9V clone of Streptococcus pneumoniae arose by large recombinational
replacements of the cpsA-pbp1a region
Microbiology 145 (Pt 8), 2023-2031 (1999)
99392464
10463168
2 (bases 1 to 2160)
Cofey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
Direct Submission
Submitted (31-MAR-1999) The Wellcome Trust Centre for the
Epidemiology of Infectious Disease, Department of Zoology,
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
3PS, UK

Location/Qualifiers
1. 2160
/organism="Streptococcus pneumoniae"
/strain="PO-273"
/serotype="14"
/db_xref="taxon:1313"
/note="Polish variant of the major penicillin-resistant
Spanish serotype 9V clone"
1. 2160
/gene="pbp1a"

```


QY	661	TAATTA	TACCTCAAGSAGTCA	TCATCA	AGTTGAGAGAAAGGCTAT	TAACCTAC	TAC	720
Db	819	TAATTA	ACCTCAAGSAGGCTCA	TCATCA	AGTAGAACAAGAAACTGGCTAT	TAACCTCT	TAAC	878
QY	721	AACGGGA	TGGATATCT	TAACCA	AAATGTAGACCAAGAAAGCTCA	AAAAACATCT	GGGATAT	780
Db	879	TACGGGAT	GGATGTTTCA	CCAAATGTAG	ACCAGAAAGCTCA	AAAAACATCT	GGGATAT	938
QY	781	TTAA	CAATACAGACCAATAC	CGTTGCTTAT	CCAGACGATGTAA	TTGGCAAGTGGCTT	CAACAT	840
Db	939	CTACAC	ATCCGATCAATATAC	GTCTCTT	TAACCTGACATGTAT	TTCAGATGCGATCT	CTAC	998
QY	841	TGTTGAT	GTGTTTCA	ACGGTAAAGTCA	TGTGCCAGCTAGGAGCAGCCAT	CACTCAAGTA	900	
Db	999	CGTGA	TATGTTTCA	AAATGTAAAGTCA	TGATCGCCAACTGGAGCTG	CAACAGATTA	1058	
QY	901	TGTTT	CCCTGGAA	TTAACCAACAGATAG	AAACCCGCACTGGGAT	TCACATATGA	960	
Db	1059	CGTTT	CAATTTGGTAC	CAACCAACGCTGG	AAACCAATCGTAC	TGGGTTTGGCTAT	9A 1118	
QY	961	ACCGAT	CAAGACATATCTCT	CGCTTGGAGTACG	GTGCTACATCTCA	ACTGCTACTAT	1020	
Db	1119	ACCAAT	CAACGATATATG	CACTGCGCATAG	ATACGTTTATGAT	TCACCTGCA	ACTAT 1178	
QY	1021	CGTT	CACATGAGCCCT	TAATACCTCG	TGGGCAAAATCTCTG	TTATTA	ACTGGGATAG 1080	
Db	1179	GGTAA	TGATATTCCTAT	TAATACATAC	CGGGAAACACAC	CGCTGTAC	AACTGGGATAG 1238	
QY	1081	GGGCT	ACTTGGCAAC	TACCTTGCA	TAATACGCTTGC	CAACAATGGCCAA	AGTCCAC 1140	
Db	1239	AGCAT	TATTTGGGTAT	TTTACTGCTG	CAATATGTCTT	CAACATCA	CGAAATGTCA 1298	
QY	1141	CGTGA	AAACCTTA	ACAAGGCGAG	CTCAACCGCCAGAC	CTTTCCTTAAT	GGTCTAG 1200	
Db	1299	CGTTG	ACACTTGTAT	TAAGTGGCTT	AGATAGAGCTTA	AAACCTTCTT	TAATGGCTTGG 1358	
QY	1201	AATG	CACTACCAAGAT	TTCTCACTCA	TAATGCAATTTCA	AGTAACACAA	CCGATTCAGA 1260	
Db	1359	TATG	CACTATCAAG	CAATGCAATTA	TATGCAAA	CGCCATTTCA	AGTAATACAA	AGAAATCTTA 1418
QY	1261	CAAAAA	TATGTGAC	CAAGTATGT	CAAAAAAGATG	CTGCTTAC	GCTTGGCAAATGG 1320	
Db	1419	TAA	CAATACGAC	CAAGTATGT	CAAAAAAGATG	CTGCTTAC	GCTTGGCAAATGG 1478	
QY	1321	TGGA	CTACTA	TAACCAATGT	ATATCCATAA	AGCTTAA	GTAGATGGGAGATGA 1380	
Db	1479	TGG	CACTACTAT	TAACCAATGT	ATATCCATAA	AGCTTAA	GTAGATGGGAGATGA 1538	
QY	1381	AGAG	TTCTCAATGT	CGAAGCTGT	GCATGAGAAAGAA	AGCCTATAT	ATATATACCGA 1440	
Db	1539	AGAG	TTCTCAATGT	CGAAGCTGT	GCATGAGAAAGAA	AGCCTATAT	ATATATACCGA 1598	
QY	1441	CATG	ATGAAAA	CACTTGT	GACTATGTG	AACTAGCAAAATG	CCCTATCTTGGCTGCC 1500	
Db	1599	CATG	ATGAAAA	CACTTGT	GACTATGTG	AACTAGCAAAATG	CCCTATCTTGGCTGCC 1658	
QY	1501	TCAG	CTGTGTAA	ACAGAACTCT	TAACCTTA	AGAGAGGAAAT	TTGAAACCAATCA 1560	
Db	1659	TCAG	CTGTGTAA	ACAGAACTCT	TAACCTTA	AGAGAGGAAAT	TTGAAACCAATCA 1718	
QY	1561	GAC	CTCTCAATTTG	TACACCTG	ATGAACTATTTG	CTGCTAT	TACAGCGTAAATATTTCAAT 1620	
Db	1719	GAC	CTCTCAATTTG	TACACCTG	ATGAACTATTTG	CTGCTAT	TACAGCGTAAATATTTCAAT 1778	
QY	1621	GGCT	GTATGAC	AGCGTAT	CTCAACGCTTG	ACACACTT	TGAGCGCATGGCCTTACGCT 1680	
Db	1779	GGCT	GTATGAC	AGCGTAT	CTCAACGCTTG	ACACACTT	TGAGCGCATGGCCTTACGCT 1838	
QY	1681	CGCT	GCAAAAGTT	TACGCGCT	TAATGATG	ACTACCTGT	GAAAGCAATCCAGAGA 1740	
Db	1839	TGC	AGCTAAATTTAT	GCTCTAT	GATGACCTAC	CTGCTG	GAAGGAAAGCAATCCAGAGA 1898	
QY	1741	TTGA	ATATAC	CAAGGCGCT	TAACAGAAATG	AGAAATTCG	ATTTAAAAATGGTGC 1800	

Db	Accession	Source	Location/Qualifiers
Db	1899	TTGGAATATACACGAGGGGCTCTACAGAAATGAGAAATTGCTATTAAATAATGCTCTCG	1958
Qy	1801	TTCTCAATGAGCAACACACCGCTCCACCAACACCCCATCACTACATGAAAGCTCAAGCTCAATC	1860
Db	1959	TTCTTACAGTGAGACTACCGCTCTCCACCAACACCCCATCACTACATGAAAGCTCAAGCTCAATC	2018
Qy	1861	ATCAGATAGTTCACACTTCCAGCTTACGCTTACGCTTCCAGCTCCAGCTCAAGCTCAAGCTCAAT	1920
Db	2019	ATCAGATAGTTCACACTTCCAGCTTACGCTTACGCTTCCAGCTCCAGCTCAAGCTCAAGCTCAAT	2078
Qy	1921	TACCAATCTCTAACAATTAATAGCAACCAATCAATCAACACCCCTGATCAACAATAATACAGAA	1980
Db	2079	TACCAATCTCTAACAATTAATAGCAACCAATCAATCAACACCCCTGATCAACAATAATACAGAA	2138
Qy	1981	TCCCTACACGACCAACCA 1999	
Db	2139	TCCCTACACGACCAACCA 2157	
RESULT 13			
LOCUS	AF139885	2160 bp DNA linear	Bct 26-AUG-1999
DEFINITION	Streptococcus pneumoniae strain PO-341 penicillin-binding protein		
ACCESSION	AF139885		
VERSION	AF139885.1	GI:5410460	
KEYWORDS			
SOURCE	Streptococcus pneumoniae.		
ORGANISM	Streptococcus pneumoniae		
REFERENCE	1 (bases 1 to 2160)		
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.		
TITLE	Serotype 14 variants of the Spanish penicillin-resistant serotype		
JOURNAL	9V clone of Streptococcus pneumoniae arose by large recombinational		
MEDLINE	99392464		
PUBLISHED	10463168		
REFERENCE	2 (bases 1 to 2160)		
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.		
TITLE	Submitted (31-MAR-1999) The Wellcome Trust Centre for the		
JOURNAL	University of Oxford, South Parks Road, Oxford, Oxfordshire OX1		
FEATURES			
source			
gene			
CDS			

GLGIDYPSMHTAHANAISNTTESNKQYGNASSSEKMAAAYAPANGSTYYIKPMYIHKVES
DGSKKEFSNVTGRAMKETETAYMTDMKTVLTJYTGGRATYJ.PMLPQAGKGTSMYDE
EVENHINNTGYAPDEMFVGTGRKYSHAVYENRLEPIVGDFFLAUKAYRSMYTY
LSEGSNEDMNIPEGLYRNGEFVKNGARSPNSPAPDOOPSTESSSSSDSSSTSOSS
STPSTNNSTTNNNNNTQOSNTTPODONNPQAPQ*

BASE COUNT 693 a 494 c 407 g 566 t

ORIGIN

Query Match 85.3%; Score 1704.6; DB 1; Length 2160;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 1815; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 1 TAAAAATACGACAAATATAAATCAATCATCTGCTGACCTGGGTTCTGAAGCCGGCTCAA 60
Db 159 TAAATATACGACAAATATAAATCAATCATCTGCTGACCTGGGTTCTGAAGCCGGCTCAA 218
QY 61 TGCCCAACTAATGATATATCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAACTAATGATATATCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACACAGGGGGATGATACCTCCGTATCCGGAGGCTTCTTCGCGAA 180
Db 279 TCGCTTCTTCGACACAGGGGGATGATACCTCCGTATCCGGAGGCTTCTTCGCGAA 338
QY 181 TCTGCAAGCAATTCCTCCAGAGTGATCAACTCTCAACCAAGATTGATTAGTTAGTAC 240
Db 339 TCTGCAAGCAATTCCTCCAGAGTGATCAACTCTCAACCAAGATTGATTAGTTAGTAC 398
QY 241 TTACTTTCAACTCTGACATTCGACACAGCTATTTCTGTAAGGCTCGAAGAGCTTGTT 300
Db 399 TTACTTTCAACTCTGACATTCGACACAGCTATTTCTGTAAGGCTCGAAGAGCTTGTT 458
QY 301 AGCGATTAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATA 360
Db 459 AGCGATTAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATA 518
QY 361 GGTCTACATGCTAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 420
Db 519 GGTCTACATGCTAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 578
QY 421 AGACCTCAATATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 480
Db 579 AGACCTCAATATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAAGCCCAAGCCCAAGCCCAAGCC 540
Db 639 AAACCAATATGACCCCTATTCACATCCAGAAAGCAAGCCCAAGCCCAAGCCCAAGCC 698
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGAAACAGTATGAGAAAGCAAGTCAATAC 600
Db 699 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGAAACAGTATGAGAAAGCAAGTCAATAC 758
QY 601 ACCAATTTCTATGAGCTACACAAAGTCTCAATTCAGCAGTAAATTCCTGCTTACATGGA 660
Db 759 ACCAATTTCTATGAGCTACACAAAGTCTCAATTCAGCAGTAAATTCCTGCTTACATGGA 818
QY 661 TAATTAACCTCAAGAGAGTCAATCAAGTGAAGAGAAAGAGAGAGAGTAACTACTAC 720
Db 819 TAATTAACCTCAAGAGAGTCAATCAAGTGAAGAGAAAGAGAGAGTAACTACTAC 878
QY 721 AACTGGGATGATGTTACACAAATGTAGACCAAGAGGCTCAAAAAACATCTGTTGGGATAT 780
Db 879 AACTGGGATGATGTTACACAAATGTAGACCAAGAGGCTCAAAAAACATCTGTTGGGATAT 938
QY 781 TTACAATACACAGCAATACGTTGGCTATCCAGAGATGAATTCGAAGTCGCTTACCAT 840
Db 939 CTACAATACGCAATACGTTGGCTATCCAGAGATGAATTCGAAGTCGCTTACCAT 998
QY 841 TGTGATGTTCTAAGCGTAAAGTCAATGTCAGAGTACAGCAGCCATCAGTCAAGTAA 900
Db 999 CGTGAATGTTCTAAGCGTAAAGTCAATGTCAGAGTACAGCAGCCATCAGTCAAGTAA 1058

QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAANAACCCGACCTGGGATCACTATGAA 960
Db 1059 CGTTTCATTTGTTACCAACCAAGCTGTGGAAACCAATGCTAGGGGTTCGTGTGAA 1118
QY 961 ACCGATCACAGTATGCTGCTGCTGAGTACAGGCTGTACGATTCACCTGCTACT 1020
Db 1119 ACCAATCACAGTATGCTGCTGCTGAGTACAGGCTGTATGATTCACCTGCTACT 1178
QY 1021 CGTTCACAGTACGCCCTTAATCTACCTCCGACAAATATCTGTTTAACTGGGATAT 1080
Db 1179 GGTATATATATTTCTTATTAATCTCCGGGAACAGCACACTGCTCACTACCTGGATAG 1238
QY 1081 GGGCTACTTGGCAACATACCTTGCATTAACGCTTCGACAAATGCGGAATGCCAGC 1140
Db 1239 AGCATATTTCCGTATATATTAATCTGCAATATGCTTCAACAAACAGAAATGTCACAGC 1298
QY 1141 CGTGAACCTTAACAAGTGTGAGCTACACCGGGCCAAAGCTTCTTAATGGTCTAGG 1200
Db 1299 CGTTGAGACTTGTGAATAAGGTGGCTAGATAGAGCTAAACCTTCTTAATGGTCTTGG 1358
QY 1201 AATGCACTAACCAAGTATTAATCTACTCAATGCAATTCAGTAAACAAACGCAATCAGA 1260
Db 1359 TATGCACTATCCAAAGCATGATTAATGCAAGCCATTTCAAGTATATACAAAGATCTAA 1418
QY 1261 CAAAAAATATGAGCAGCAAGTATGAAAAGATGCTGCTGCTTACGCTGCTTTCGCAATGG 1320
Db 1419 TAAACCAATACGCAAGCAAGTATGAAAAAATGGCTGCTTATCTGCTTTCGCAAAATGG 1478
QY 1321 TGGAACTTACTAATAACCAATGATATTCATTAAGTGTCTTTAGTATGGAGCTGAAA 1380
Db 1479 TGGCACTTACTAATAACCAATGATATTCATTAAGTGTCTTTAGTATGGAGCTGAAA 1538
QY 1381 AGAGTTCTCTAATGTCGGAACCTGTCGATGAGAAAGCAAGCCTATATATGAGACGA 1440
Db 1539 AGAGTTCTCTAATGTCGGAACCTGTCGATGAGAAAGCAAGCCTATATATGAGACGA 1598
QY 1441 CATGATGAAAAACGCTTGAATTAATGAACTGCAAGTGAATGCTTATCTGCTGGCTCC 1500
Db 1599 CATGATGAAAAACGCTTGAATTAATGAACTGCAAGTGAATGCTTATCTGCTGGCTCC 1658
QY 1501 TCAGGCTGTGTAACAGGAACCTTAATCTATACAGACAGAGAAATGAAAACCAATCAA 1560
Db 1659 TCAAGCTGTGTAACAGGAACCTTAATCTATACAGATGAGAGAAATGAAAACCAATCAA 1718
QY 1561 GACCTCTCAATTTGTAGACCTGATGAACCTATTTGCTGGCTATACGCTTAATTTCAAT 1620
Db 1719 GAACACTGCTATGATAGTCCAGATGAATGTTGTGGTATATCTGTAATGTTCTAT 1778
QY 1621 GGTGTATGACAGGCTATTTCAACGCTGTGACACCACTGTGAGCAATGGCTTACGCT 1680
Db 1779 GGTGTATGACAGGCTATTTCAACGCTGTGAACTGCTATGCTGAGATGGTTTCTTAGT 1838
QY 1681 CGCTGCCAAAGTTTACGCGCTATGATGACCTACTGCTGTAAGGAAGCAATCCAGAGA 1740
Db 1839 TGCAGCTTAAAGTTTATGCTTATGATGACCTACTGCTGTAAGGAAGCAATCCAGAGA 1898
QY 1741 TTGGAATATACCAAGGGGCTCTACAGAAATGAGAAATTTGTAATTTAAAAATGCTGCTG 1800
Db 1899 TTGGAATATACCAAGGGGCTCTACAGAAATGAGAAATTTGTAATTTAAAAATGCTGCTG 1958
QY 1801 TTCTAGCGTGAACCTACCTGCTCCACAACAACCCCATCAACCTGAAGTTCAAGCTCATC 1860
Db 1959 TTCTAGCGTGAACCTACCTGCTCCACAACAACCCCATCAACCTGAAGTTCAAGCTCATC 2018
QY 1861 ATCAGATAGTCAACTTACAGCTAGCTACCTCAACCAAGCAAGCAATATATGTAGAC 1920
Db 2019 ATCAGATAGTCAACTTACAGCTAGCTACCTCAACCAAGCAAGCAATATATGTAGAC 2078
QY 1921 TACCAATCTTAACCAATATATAGCAACCAATCAACCAATCAACCAATCAACCAATCA 1980
Db 2079 TACCAATCTTAACCAATATATAGCAACCAATCAACCAATCAACCAATCAACCAATCA 2138
QY 1981 TCTTCAACCAAGCACAACCA 1999

|||||
Db 2139 TCCTCAGCAGCACAACCA 2157
RESULT 14
AF139886 2160 bp DNA linear BCT 26-AUG-1999
LOCUS Streptococcus pneumoniae strain FO-342 penicillin-binding protein
1a (pbpla) gene, complete cds.
ACCESSION AF139886
VERSION AF139886.1 GI:5410462
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE Serotype 14 variants of the Spanish penicillin-resistant serotype
9v clone of Streptococcus pneumoniae arose by large recombinational
replacements of the cpsA-pbpla region
JOURNAL Microbiology 145 (Pt 8), 2023-2031 (1999)
MEDLINE 99392464
PUBMED 10463168
REFERENCE 2 (bases 1 to 2160)
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-1999) The Wellcome Trust Centre for the
Epidemiology of Infectious Disease, Department of Zoology,
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
3PS, UK
FEATURES
source Location/Qualifiers
1. 2160
/organism="Streptococcus pneumoniae"
/strain="FO-342"
/serotype="14"
/db_xref="taxon:1313"
/note="Polish variant of the major penicillin-resistant
Spanish serotype 9v clone"
1. 2160
/gene="pbpla"
1. 2160
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=1
/product="penicillin-binding protein 1a"
/protein_id="AAD43069.1"
/db_xref="GI:5410463"
/translation="MNKPTILRLIKYLSIFSLVIAIVLGGVFEVYVSKAPSLSE
SKIVATSSKTYDNKNQLIADLSRRRYNAANDPVLKATVSIEDHFRPDHIGTD
SIRLGAFLRLQNSLDGSLTFOOLIKLYFSTSDQTSRAQDAMLAIOLECK
ATKOEILYIVINKYVMSNGYGMQIAONYGKDLNNLSLQALALAMPQAPNDYD
YSHPEAQRRLNLYLSEKNGOYISAEOYEAIVPPIIDGQSLKLSANYPAYMNDYL
KEVINOVFOETGYNLITGMVYTVNVDQKHLMDIYNSOQYVSPDDDOVASTAV
DVSNGKVIADLGARHOASNGTMOAETNRDMGSAKPIIDVAPALAEVYVSTAT
MVNDIPYRPTSTPYVMDPARENTILQIALDQSRVTAVENTLNKGLDRKFFLN
GLGIDPEMHYANALISNTTESNKQYGAASSSEKMAAIVAFANGGTYIPMTIHKVF
DGSKEEFSNVGTRAKETAYMTDMKTVLTGTGRALYLPWLEQAKGTSTNTDE
EVENHIKNTGYVADPEMFVYTRKYSMAVMGYSNRLPIVGDGLVAKYRSMY
LSEGSNPEDMNIPEGIVRNGEVEFKGARSSTMSPAPOQPTSESSSSSDSTSS
STPSTNSTTNNNTNNTPOQSNTPDOQNQNPQAP"
BASE COUNT 693 a 494 c 407 g 566 t
ORIGIN
Query Match 85.3%; Score 1704.6; DB 1; Length 2160;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 1815; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 61 TCCCCAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTATCGAAGACCA 120
Db 219 TCCCCAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACACAGAGGGGATGATACATCCGTATCTGGAGCTTTCTGGCAA 180
Db 279 TCGCTTCTTCGACACAGAGGGGATGATATTCATCCGTATCTGGAGCTTTCTGGCAA 338
QY 181 TCTGCAAGCAATTCCTCTCAGAGTGATGATCACTCCACCCAAAGTTGATTAAGTGAC 240
Db 339 TCTGCAAGCAATTCCTCTCAGAGTGATGATCACTCCACCCAAAGTTGATTAAGTGAC 398
QY 241 TTACTTTTCACTTCGACTTCCGACAGACTATTTCTCTAAGGCTCAGAGACTTGGTT 300
Db 399 TTACTTTTCACTTCGACTTCCGACAGACTATTTCTCTAAGGCTCAGAGACTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACTTACTATATAATA 360
Db 459 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACTTACTATATAATA 518
QY 361 GGTCTACATGTCTAATGGAACTATGGAATGCAGACAGAGCTCAAACTACTATGTAA 420
Db 519 GGTCTACATGTCTAATGGAACTATGGAATGCAGACAGAGCTCAAACTACTATGTAA 578
QY 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGGTGGGAATGCTCAGGACAC 480
Db 579 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGGTGGGAATGCTCAGGACAC 638
QY 481 AAACCAATATGACCCCTATTTCAATCCAGACAGCCCAAGACCCGCAAACTTGGTCT 540
Db 639 AAACCAATATGACCCCTATTTCAATCCAGACAGCCCAAGACCCGCAAACTTGGTCT 698
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACGTTAGAAAGCACTCAATAC 600
Db 699 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACGTTAGAAAGCACTCAATAC 758
QY 601 ACCAATTAATGATGAGCTCAAGAGTCTCAAAAGTCAAGCAAGTAATTCCTGCTTACATGA 660
Db 759 ACCAATTAATGATGAGCTCAAGAGTCTCAAAAGTCAAGCAAGTAATTCCTGCTTACATGA 818
QY 661 TAATTACCTCAAGAGAGTCAATCAATCAAGTGAAGAAAGCAAGGCTAATACCTACAC 720
Db 819 TAATTACCTCAAGAGAGTCAATCAATCAAGTGAAGAAAGCAAGGCTAATACCTACAC 878
QY 721 AACTGGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
Db 879 TACTGGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 938
QY 781 TTACAATACAGAGCAATAGCTTGGCTATCCAGACGATGAATTCGAAGTGGCTTACCAT 840
Db 939 TTACAATACAGAGCAATAGCTTGGCTATCCAGACGATGAATTCGAAGTGGCTTACCAT 998
QY 841 TGTGATGCTTTTCAACGTTAAGTCAATTCCTCAGAGTACAGACAGCCATCAAGTAA 900
Db 999 GGTGATGCTTTTCAACGTTAAGTCAATTCCTCAGAGTACAGACAGCCATCAAGTAA 1058
QY 901 TGTTCCTTTCGAAATTAACCAAGTAGAACAACCGGACTGGGGATCACTATGAA 960
Db 1059 GGTTCCTTTCGAAATTAACCAAGTAGAACAACCGGACTGGGGATCACTATGAA 1118
QY 961 ACCGATACAGAGTATGCTGCTGCTTGGATGAGGCTGCTAGATCAAGTCAATCAT 1020
Db 1119 ACCGATACAGAGTATGCTGCTGCTTGGATGAGGCTGCTAGATCAAGTCAATCAT 1178
QY 1021 CGTTCAGATGAGCCCTATTAATCACTTGGAGCAAAATACCTCTGTTTAACTGGGATAG 1080
Db 1179 GGTTCAGATGAGCCCTATTAATCACTTGGAGCAAAATACCTCTGTTTAACTGGGATAG 1238
QY 1081 GGGCTACTTGGCAACATCACTTGGCAATAGCCCTGCAACAATGCGCAAGCTCCAGC 1140
Db 1239 ACCATATTTGGTAAATATATACCTGCAATATGCTCTTCAACAATCAAGAAATGTCACAGC 1298

Qy	1141	CGTGAACCTCTAAACAAAGTCGAGCTCAACCGCCCAAGACTTTCCTAATAGTCTAGG	1200
Db	1299	CGTTGAGCTTTGAATTAAGTCGGCTCTAGATAGAGCTAAACCTTCCTTAATAGTCTTGG	1358
Qy	1201	AATGCACTACCAAGATATTCTACTACTCAATGGCCATTTTCAAGTAAACACACGATCAGA	1260
Db	1359	TATGACACTATCCACAGCATATGATTAAGCAACGCCATTTCAAGTATACACAGAAATCGAA	1418
Qy	1261	CAAAAATATGAGCAAGAGTAAAGATGGAATGGCTGCTGCTTACGCTTGGCTTGAATGG	1320
Db	1419	TAAACAAATACGAGCAAGTATGAAAAATGGCTCTGCTTATGCTGCCCTTTGCCAAATGG	1478
Qy	1321	TGGAACCTACTATAAACCAATGTATATCCATAAAGTCGCTTACTAGTGGAGTGAATA	1380
Db	1479	TGGCACTTACTATAAACCAATGTATATCCATAAAGTCGCTTACTAGTGGAGTGAATA	1538
Qy	1381	AGAGTTCTCTAATGTGGAACTCGTCCATGAGAAACGACAGCCTATATGATGACCGA	1440
Db	1539	AGAGTTCTCTAATGTGGAACTCGTCCATGAGAAACGACAGCCTATATGATGACCGA	1598
Qy	1441	CATGATGAAAAAGTCTGACTTATGGAAGTGAAGCAATGCTTATGCTTGGCTGCC	1500
Db	1599	CATGATGAAAAAGTCTGACTTATGGAAGTGAAGCAATGCTTATGCTTGGCTGCC	1658
Qy	1501	TCAGGCTGTAAACAGAGAACCTTACTATACAGACGAGAAATTTGAAAACACATCAA	1560
Db	1659	TCAAGCTGTAAACAGAGAACCTTACTATACAGATGAGAAAGTTGAAAACACATCAA	1718
Qy	1561	CACCTCTCAATTTTGACAGCCTGATGAACATATTTCCTGCTATACGCGTAATATTCAAT	1620
Db	1719	GAACTCTGCGATGTAGCTCAGATGAATGTTTGTGGTATACTCGTAAGTATTCTAT	1778
Qy	1621	GGCTTATGAGCAGGCTTATCTTAACCGTCTGACACCACTTGGAGCAATGGCCTTACGGT	1680
Db	1779	GGCTTATGAGCAGGCTTATCTTAACCGTCTGATCTTAACTCTTATCGTTGAGATGGTTTCTTACT	1838
Qy	1661	CGCTGCAAAAGTTTACCGCTCTATGATGACCTTACCTGCTCTGAAGAGCAATCCAGAGA	1740
Db	1839	TGCAGCTTAAAGTTTATCCGCTCATATGACCTTACCTGCTCTGAAGAGCAATCCAGAGA	1898
Qy	1741	TTTGAATATACCAAGGGGCTCTACAGAAATGGAATTTTCGATTTTAAAAATGGTCTCG	1800
Db	1899	TTTGAATATACCAAGGGGCTCTACAGAAATGGAATTTTCGATTTTAAAAATGGTCTCG	1958
Qy	1801	TTTCTGCGGAATCAGCTGCTCCACAAACCCCATCACTGAAGTTCAAGCTCATC	1860
Db	1959	TTTCTGCGGAGCTCAGCTGCTCCACAAACCCCATCACTGAAGTTCAAGCTCATC	2018
Qy	1861	ATCAGATAGTTCAACTTACAGTCTAGCTCAACCACTCCAAAGCACAATAATATAGTAGAC	1920
Db	2019	ATCAGATAGTTCAACTTACAGTCTAGCTCAACCACTCCAAAGCACAATAATATAGTAGAC	2078
Qy	1921	TACCAATCTTAACATATATAGCAACATTCAAATTCAAACCCCTGATCAACAAATACAGA	1980
Db	2079	TACCAATCTTAACATATATAGCAACATTCAAATTCAAACCCCTGATCAACAAATACAGA	2138
Qy	1981	TCCTCAACGACACACACA 1999	
Db	2139	TCCTCAACGACACACACA 2157	

RESULT 15
AF139887 2160 bp DNA linear BCT 26-AUG-1999
LOCUS Streptococcus pneumoniae strain M134 penicillin-binding protein 1a
DEFINITION (pbpla) gene, complete cds.
ACCESSION AF139887
VERSION AF139887.1 GI:5410464
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.

REFERENCE	(bases 1 to 2160)	
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.	
TITLE	Serotype 14 variants of the Spanish penicillin-resistant serotype 9v clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA-pbpA region	
JOURNAL	Microbiology 145 (Pt 8), 2023-2031 (1999)	
MEDLINE	99392464	
PUBMED	10463168	
REFERENCE	2 (bases 1 to 2160)	
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.	
TITLE	Direct Submission	
JOURNAL	Submitted (31-MAR-1999) The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK	
FEATURES	Location/Qualifiers	
source	1..2160	
	/organism="Streptococcus pneumoniae"	
	/strain="M134"	
	/serotype="14"	
	/db_xref="taxon:1313"	
	/note="Danish variant of the major penicillin-resistant Spanish serotype 9v clone"	
gene	1..2160	
	/gene="pbpla"	
CDS	1..2160	
	/gene="pbpla"	
	/function="peptidoglycan transglycosylase-transpeptidase"	
	/codon_start=1	
	/transl_table=1	
	/product="penicillin-binding protein 1a"	
	/protein_id="AAD3070.1"	
	/db_xref="GI:5410465"	
	/translation="MNKPTILRIKIVSISFLSIVTAIVLGGVFVYVSKAPLSSE SKLVATSSKRIYDKNLIADLGSERRVNAQANDIPDVKAIYSIEDHFRFORGD SIRLGAFLRLQSNLSLQGGSTLLOQLIKLTSTSDQTSKAOEAMWLALDEK ATKOELLITLYINKVYMSNGNGVQMOYAAQNTYTGKDLNLSLPOLALLGAMPQADYD YSHPEAODRNLVLSEMKNGYISAQYEKAVNTPPTDGLQSLKASNPAYVDNL KEVINEOYEETGYLLTTGMDVYTNVOEAKHLMIDLYNSDOVSYDDDLQVASTV DVSNGKVIADLGARHQAQSNVSGFNQVAFERPMDSAMKPTIDVAPLIEGVDSSTAT MVDNIPYNIQGTSPYVNPMDRAVEGNTTIOYALDOSSNVAVETLNKVGIDRAKTFIN GLGIDYPMHRTANALISNTTESNKQYASSSSKMAAATAAFRANGSTYIKPMYIRKVS DGSKKEFSNVGTRAMKEPTAYMMTDMKLVLYTGEGVALPWLQPGKGTGTSNYDE EVENHINKNTGVAPADEMFVGYTRKYSAAVNTVYSGNRLPTVGDGLVAARVRSMTY LSEDTHPEDMTMPDGLFRNGEFVFKNARSWNPAQOPPESTRESSSSSDSSSTOSS STPSTNNSTTTPNNNTTQQSNTPPDQONQNPQAPQ"	
BASE COUNT	696 a	495 c 402 g 567 t
ORIGIN		
Query Match	83.9%;	Score 1677.4; DB 1; Length 2160;
Best Local Similarity	89.9%;	Pred. No. 0;
Matches 1798; Conservative	0; Mismatches 201; Indels 0; Gaps 0;	
QY	1 TAAATCTAGACATAAAAAATCAACTATGCTGACTTGGGTTCTGAACCGCGGTCAA	60
Db	159 TAAATCTAGACATAAAAAATCAACTATGCTGACTTGGGTTCTGAACCGCGGTCAA	218
QY	61 TGCCCAAGCTATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACCA	120
Db	219 TGCCCAAGCTATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACCA	278
QY	121 TCGCTTCTTGACCAAGGGGATTTGATCAGTCGATCCGTTGGGAGCTTCTTGCGCAA	180
Db	279 TCGCTTCTTGACCAAGGGGATTTGATCAGTCGATCCGTTGGGAGCTTCTTGCGCAA	338
QY	181 TCTGCAAAAGTAATTCCTCCAAAGTGATCAACTCCACCAAGTGTATTAAGTTGAC	240
Db	339 TCTGCAAAAGTAATTCCTCCAAAGTGATCAACTCCACCAAGTGTATTAAGTTGAC	398
QY	241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTGAAGGCTCAGGAAGCTTGGTT	300
Db	399 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTGAAGGCTCAGGAAGCTTGGTT	458

QY 301 AGCGATTCAGTTAGAACAAAAGCAACGAAGAAATCTTGACCTACTATATAATAA 360
 Db 459 AGCGATTCAGTTAGAACAAAAGCAACGAAGAAATCTTGACCTACTATATAATAA 518
 QY 361 GGTCTACATGCTTAATGGGAACTATGGAATGAGAGAGAGCTCAAACTACTATGTA 420
 Db 519 GGTCTACATGCTTAATGGGAACTATGGAATGAGAGAGAGCTCAAACTACTATGTA 578
 QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTCAGCACC 480
 Db 579 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTCAGCACC 638
 QY 481 AAACCAATATGACCCCTATTCATCCAGAGAGCCGCAAGCCGCGCAAACTGGTCTT 540
 Db 639 AAACCAATATGACCCCTATTCATCCAGAGAGCCGCAAGCCGCGCAAACTGGTCTT 698
 QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGATATGAGAAAGCAGTCAATAC 600
 Db 699 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGATATGAGAAAGCAGTCAATAC 758
 QY 601 ACCAATTAATGAGCTACAGTCAAAAGTCTCAATCAGCAATTAATACCTGCTTACATGA 660
 Db 759 ACCAATTAATGAGCTACAGTCAAAAGTCTCAATCAGCAATTAATACCTGCTTACATGA 818
 QY 661 TAATTAACCTCAAGGAGTCAATCAATCAATGAGAAAGAGAGCTATTAACCTACTCAC 720
 Db 819 TAATTAACCTCAAGGAGTCAATCAATCAATGAGAAAGAGAGCTATTAACCTACTCAC 878
 QY 721 AACTGGATGATGCTCTACAAATGTAAGCAGAGAGCTCAAAACATCTGTGGATAT 780
 Db 879 TACTGGATGATGCTCTACAAATGTAAGCAGAGAGCTCAAAACATCTGTGGATAT 938
 QY 781 TTACAATACAGAGATACCTGCTATCCAGAGATGAATTGCAAGTCTTCAATCAT 840
 Db 939 TTACAATACAGAGATACCTGCTATCCAGAGATGAATTGCAAGTCTTCAATCAT 998
 QY 841 TTCTGATGTTCTAAGGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Db 999 TTCTGATGTTCTAAGGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1058
 QY 901 TCTTTCCTTCGGAATTAACCAAGCAGTGAACAAACCCGAGTGGGATCAACTATGA 960
 Db 1059 CGTTTCATTTGGTACCAACCAAGCTGTGAACCAATCTGTGAGGCTTCTGCTATGA 1118
 QY 961 ACCGATCAGAGTATGCTGCTGCTTGGAGTACGCTGCTACAGTCACTGCTACTAT 1020
 Db 1119 ACCGATCAGAGTATGCTGCTGCTTGGAGTACGCTGCTTATGATTCACAGTCACTAT 1178
 QY 1021 CGTTACGATGAGCCCTATTAATACCTGGGACCAATACTCCTGTTTATTAATGGGATAG 1080
 Db 1179 GGTTAATGATTAATCTTAATTAATCTCGGGAAACGACACCTGCTTACAACTGGGATAG 1238
 QY 1081 GGGCTACTTTGGCAACATCAGCTTGAATACGCTGCAACATGCGAAACGCTGCCAGC 1140
 Db 1239 AGCATATTTGGTAATTAATTAATCTGCAATATGCTCTTCAACAATCAGAAATGTCAAGC 1298
 QY 1141 CGTGAACCTTAACCAAGGCTGAGCTCAACGCGCGCAAGACTTTCCTAATGCTTAGG 1200
 Db 1299 CGTTGAGACTTGAATTAAGTGGTGTAGATAGAGCTAAACCTTCCTTAATGCTTAGG 1358
 QY 1201 AATGCACTACCAAGATTAATCACTACCAATGCAATTTCAAGTAAACCAACGAAATCAGA 1260
 Db 1359 TATGCACTATTCAGATCAATTAATGCAACGCAATTTCAAGTAAATCAACGAAATCTAA 1418
 QY 1261 CAAAAAATATGAGCAAGTATGAAAAAGATGGTGGTCTTACGCTGCTTGGCAATAGG 1320
 Db 1419 TAAACATATGAGCAAGTATGAAAAAGATGGTGGTCTTACGCTGCTTGGCAATAGG 1478
 QY 1321 TGGAACTTACTATTAACCAATGATATATCAATAAGTCTTATAGTGGAGTGAAGA 1380
 Db 1479 TGGCACTTACTATTAACCAATGATATATCAATAAGTCTTATAGTGGAGTGAAGA 1538

QY 1381 AGAGTTCTTAATGTCGGAACCTGTCACATGAGAGAAAGACAGCTATATGATGACCGA 1440
 Db 1539 AGAGTTCTTAATGTCGGAACCTGTCACATGAGAGAAAGACAGCTATATGATGACCGA 1598
 QY 1441 CATGATGAAAAAGCTCTTGAATTAATGGAACCTGACGAAAAATGCCATCTTGGTGGCTCC 1500
 Db 1599 CATGATGAAAAAGCTCTTGAATTAATGGAACCTGACGAAAAATGCCATCTTGGTGGCTCC 1658
 QY 1501 TCAAGCTGGTAAAAAGCAAGCACTTAATTAATGAGAGAGAGAAATGAAAAACCAATCAA 1560
 Db 1659 TCAAGCTGGTAAAAAGCAAGCACTTAATTAATGAGAGAGAGAAATGAAAAACCAATCAA 1718
 QY 1561 GACCTCTCAATTTGTAGCACTGATGACATTAATTTGGTGGCTATAGCGTAATATTCATAT 1620
 Db 1719 GACCTCTCAATTTGTAGCACTGATGACATTAATTTGGTGGCTATAGCGTAATATTCATAT 1778
 QY 1621 GGGCTATGAGAGAGCTATTAATCAACGCTGACAGCACTTGAAGCAATGGCTTACGGT 1680
 Db 1779 GGGCTATGAGAGAGCTATTAATCAACGCTGACAGCACTTGAAGCAATGGCTTACGGT 1838
 QY 1681 CGCTGCCAAAGTTTACCGCTTATGATGACCTTACCTGCTGAAGAGAGCAATGCAAGA 1740
 Db 1839 TGCAGCTAAAGTTTATCGCTCAATGATTAACGATATCTATGAGAGATACATCAGAGA 1898
 QY 1741 TTGGAATATACAGAGAGGGGCTACAGAAATGAGAAATGGTATTAATAATGGTCTCG 1800
 Db 1899 CTGGAGATGACAGAGAGCTTTTCAGAAATGAGAAATGGTATTTAAATATGGTCTCG 1958
 QY 1801 TTCTACGTGGAACCTACCTGCTTCACAGCAACCCCATCACTGAAGTTCAGAGCTATC 1860
 Db 1959 TTCTACGTGGAACCTACCTGCTTCACAGCAACCCCATCACTGAAGTTCAGAGCTATC 2018
 QY 1861 ATCAGATGTTCACTTACAGTCTAGCTCAACCACTCCAGAGCAAAATATATAGTACGAC 1920
 Db 2019 ATCAGATGTTCACTTACAGTCTAGCTCAACCACTCCAGAGCAAAATATATAGTACGAC 2078
 QY 1921 TACCAATCTTAACCAATTAATGCAACCAATTAATTAACCCCTGATCAACCAAAATGAGA 1980
 Db 2079 TACCAATCTTAACCAATTAATGCAACCAATTAATTAACCCCTGATCAACCAAAATGAGA 2138
 QY 1981 TCCTCAACGACCAACCA 1999
 Db 2139 TCCTCAACGACCAACCA 2157

Search completed: August 11, 2002, 03:11:46
 Job time: 28792 sec

•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2002, 01:12:49 ; Search time 325.81 seconds
(without alignments)
10534.090 Million cell updates/sec

Title: US-08-961-083-1
1999
Perfect score: 1 TAAATCTACGACATATAA.....ATCCTCAACGACACACCA 1999
Sequence:

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_032802:*

- 1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	19	AAV27323
2	1999	100.0	10711	19	AAV52278
3	1981.4	99.1	2048	17	AAT08027
4	1981.4	99.1	2160	22	AAH02055
5	1981.4	99.1	2160	23	AAS55688
6	1677.4	83.9	2160	22	AAH02056
7	1642.2	82.2	2160	22	AAH01178
8	1621.4	81.1	2160	22	AAH02054
9	1607	80.4	2157	22	AAH02147

10	1252	62.6	1260	20	AAZ35952	Streptococcus pneu
11	1248.8	-62.5	1260	20	AAZ35950	Streptococcus pneu
12	1248.8	62.5	1260	20	AAZ35951	Streptococcus pneu
13	1248.8	62.5	1260	20	AAZ35953	Streptococcus pneu
14	1247.4	62.4	1249	22	AAH02067	Streptococcus pneu
15	1245.6	62.3	1260	20	AAZ35939	Streptococcus pneu
16	1202.4	60.2	1260	20	AAZ35948	Streptococcus pneu
17	1191.2	59.6	1260	20	AAZ35949	Streptococcus pneu
18	1183.2	59.2	1260	20	AAZ35947	Streptococcus pneu
19	1167.2	58.4	1260	20	AAZ35946	Streptococcus pneu
20	1162.4	58.1	1260	20	AAZ35943	Streptococcus pneu
21	1160.8	58.1	1260	20	AAZ35944	Streptococcus pneu
22	1122.2	56.1	1199	22	AAH01023	Streptococcus pneu
23	1108	55.4	1260	20	AAZ35945	Streptococcus pneu
24	1073.8	53.7	1229	22	AAH01026	Streptococcus pneu
25	1053.6	52.7	1212	22	AAH02063	Streptococcus pneu
26	949.6	47.5	1260	20	AAZ35942	Streptococcus pneu
27	944.2	47.2	960	16	AAQ83241	Streptococcus pneu
28	938.4	46.9	1260	20	AAZ35940	Streptococcus pneu
29	938.4	46.9	1260	20	AAZ35941	Streptococcus pneu
30	897.2	44.9	1222	22	AAH01025	Streptococcus pneu
31	895.2	44.8	1220	22	AAH01022	Streptococcus pneu
32	890.2	44.5	1223	22	AAH01017	Streptococcus pneu
33	885.2	44.3	1218	22	AAH01016	Streptococcus pneu
34	881.2	44.1	1214	22	AAH01018	Streptococcus pneu
35	879.2	44.0	1212	22	AAH01013	Streptococcus pneu
36	875.4	43.8	1213	22	AAH01015	Streptococcus pneu
37	874.4	43.7	1212	22	AAH01014	Streptococcus pneu
38	874.2	43.7	1207	22	AAH01020	Streptococcus pneu
39	873.4	43.7	1211	22	AAH01024	Streptococcus pneu
40	867.6	43.4	930	22	AAH02060	Streptococcus pneu
41	863	43.2	1223	22	AAH01019	Streptococcus pneu
42	857.6	42.9	1216	22	AAH02069	Streptococcus pneu
43	853.8	42.7	1225	22	AAH01027	Streptococcus pneu
44	841.4	42.1	1195	22	AAH02059	Streptococcus pneu
45	834.6	41.8	1201	22	AAH01021	Streptococcus pneu

ALIGNMENTS

RESULT 1
AAV27323
ID AAV27323 standard; DNA; 1999 BP.
XX AAV27323:
XX
XX
XX 02-OCT-1998 (first entry)
DE Streptococcus pneumoniae SP001 nucleotide.
XX
XX Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
KW detection; pneumonia; otitis media; meningitis; ss.
XX
XX Streptococcus pneumoniae.
OS
FH key Location/Qualifiers
FT CDS 2..1999 /*tag= a
FT /*product= "SP001"
FT /*note= "no stop codon given"
XX
XX
XX W09818930-A2.
XX
XX
XX 07-MAY-1998.
XX
XX 30-OCT-1997; 97WO-US19422.
XX
XX 31-OCT-1996; 96US-0029960.
XX
XX (HDMA-) HDMA GENOME SCI INC.
XX
XX Chol GH, Hromockyj A, Johnson LS, Kunsch CA;

Db 1681 cgcgtccaaagttaacgcgtatgatgacctactgctgtctgaaggagcaatccagaaga 1740
 Qy 1741 TTGGAATATACGAGGGGGCTCTACGAATATGAGAAATTCGATTTTAAATATGTCG 1800
 Db 1741 ttggaatatacagaggggctctacgaatattgagaattcgtatttaaaatggtctcg 1800
 Qy 1801 TTTCTGAGGAAGTACCTGCTCCACCAACACCCCATCAATGGAAGTTCAGGTCATC 1860
 Db 1801 ttctacgtggaactcaccgtcccaacaaccccccaactgaagtlcaagctcacc 1860
 Qy 1861 ATCAGATGTTCAACTTCACTGCTAGCTCAACCTCCACGACCAATATATGTCAGC 1920
 Db 1861 atcagatagttcaacttcaactcgtctagctcaaccctcaagcaacaataatagtcagac 1920
 Qy 1921 TACCAATCTTCAATTAATAGCAACATCAATCAACCCCTGATCAACAAATCAGAA 1980
 Db 1921 taccatctcacaataatacgcacaacatcaataaccctgltcaacaataatcagaa 1980
 Qy 1981 TCCCAACGACGACCAACCA 1999
 Db 1981 tcccaaccgacacaacca 1999
 RESULT 2
 AAV52278/C
 ID AAV52278 standard; DNA; 10711 BP.
 AC AAV52278;
 DT 23-OCT-1998 (first entry)
 DE Streptococcus pneumoniae genome fragment SEQ ID NO:145.
 XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
 KW computer readable medium; vaccine; pharmaceutical composition; ds.
 XX Streptococcus pneumoniae.
 OS Streptococcus pneumoniae.
 XX W09818931-A2.
 PD 07-MAY-1998.
 XX 30-OCT-1997; 97WO-US19588.
 PR 31-OCT-1996; 96US-0029960.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
 PI Kunsch CA, Rosen CA;
 DR WPI; 1998-272225/24.
 XX Computer-readable medium with recorded Streptococcus pneumoniae
 PT polynucleotide sequences - useful in diagnostic kits and assays, and
 PT pharmaceutical compositions and vaccines for Streptococcus
 PT pneumoniae
 PS Claim 1; Page 972-978; 1409pp; English.
 XX The present invention describes a computer readable medium which has
 CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
 CC recorded on it, or a representative fragment or a sequence at least 95%
 CC identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in
 CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
 CC Streptococcus pneumoniae. The present invention also describes an
 CC isolated nucleic acid molecule encoding a homologue of any of the
 CC fragments of the S. pneumoniae genome (SEQ ID NO:1 to 391) where the
 CC nucleic acid molecule is produced by a process comprising: (a) screening
 CC a genomic DNA library using as a probe a target sequence defined by any
 CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
 CC library which contain sequences that hybridise to the target sequence and
 CC isolating the nucleic acid molecules from the members; or (b) isolating

CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
 CC molecules whose nucleotide sequence is homologous to amplification
 CC primers derived from the fragment of the S. pneumoniae genome to prime
 CC the amplification and isolating the amplified sequences. The computer
 CC readable medium can be used in a computer-based system for identifying
 CC fragments of the S. pneumoniae genome of commercial importance, or
 CC expression modulating fragments of the S. pneumoniae genome. Products
 CC from the present invention can be used in diagnosis kits and assays, and
 CC pharmaceutical compositions and vaccines for S. pneumoniae.
 XX
 SQ Sequence 10711 BP; 3280 A; 2056 C; 2412 G; 2963 T; 0 other;
 Query Match 100.0%; Score 1999; DB 19; Length 10711;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TAAATCTACGACATPAAAAATCACTCATTCGCTGCTGCTGGAAGCGCGGTCA 60
 Db 9767 TAAATCTACGACATPAAAAATCACTCATTCGCTGCTGCTGGAAGCGCGGTCA 9708
 Qy 61 TCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTCTATGGAAGACA 120
 Db 9707 TCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTCTATGGAAGACA 9648
 Qy 121 TCGCTTCTTCGACACAGGGGGATGATACATCCGATTCCTGGGAGCTTCTTGCGCAA 180
 Db 9647 TCGCTTCTTCGACACAGGGGGATGATACATCCGATTCCTGGGAGCTTCTTGCGCAA 9588
 Qy 181 TCGCAAGCAATTCCTCCAGAGTGATGATCACTCCACCAAGTTGATTAAGTTGAC 240
 Db 9587 TCTGCAAGCAATTCCTCCAGAGTGATGATCACTCCACCAAGTTGATTAAGTTGAC 9528
 Qy 241 TTACTTTTCAACTTGCAGCTTCGACAGACATATTTCTGTAAGGCTCAGAAAGCTTGGTT 300
 Db 9527 TTACTTTTCAACTTGCAGCTTCGACAGACATATTTCTGTAAGGCTCAGAAAGCTTGGTT 9468
 Qy 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACTCTATATATTA 360
 Db 9467 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACTCTATATATTA 9408
 Qy 361 GGTCTACATGCTAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 420
 Db 9407 GGTCTACATGCTAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 9348
 Qy 421 AGACCTCATATATTTAAGTTTACCTAGTTTACCTTGTGCTGGAATGCTCAGGACAC 480
 Db 9347 AGACCTCATATATTTAAGTTTACCTAGTTTACCTTGTGCTGGAATGCTCAGGACAC 9288
 Qy 481 AAACCAATATGACCCCTATTCATCAGACAGACAGCCCAAGCCGCAAACTTGCTTT 540
 Db 9287 AAACCAATATGACCCCTATTCATCAGACAGACAGCCCAAGCCGCAAACTTGCTTT 9228
 Qy 541 ATCTGAATGAATAATCAAGGCTACATCTGCTGAACGTTGAAAGCAAGCAATAC 600
 Db 9227 ATCTGAATGAATAATCAAGGCTACATCTGCTGAACGTTGAAAGCAAGCAATAC 9168
 Qy 601 ACCAATTCATGATGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGA 660
 Db 9167 ACCAATTCATGATGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGA 9108
 Qy 661 TAATTAACCTGCAAGGATCATCATCAATGTTGAAGAAGCAAGGCTATTAACCTACTAC 720
 Db 9107 TAATTAACCTGCAAGGATCATCATCAATGTTGAAGAAGCAAGGCTATTAACCTACTAC 9048
 Qy 721 AACTGGGATGATGCTTACCAAAATGATGACCAAGAGCTCAAAAACATCTGGGATAT 780
 Db 9047 AACTGGGATGATGCTTACCAAAATGATGACCAAGAGCTCAAAAACATCTGGGATAT 8988
 Qy 781 TTAACATACAGACGATATAGCTTATCCATCAGACGATGAATTCAGATGCTTCAACAT 840
 Db 8987 TTAACATACAGACGATATAGCTTATCCATCAGACGATGAATTCAGATGCTTCAACAT 8928

QY 841 TGTGATGTTTCTAAGCTAAAGTCATTCGCCAGCTAGGAGCGCCATCAGTCAGTAA 900
 |||||||
 Db 8927 TGTGATGTTTCTAAGCTAAAGTCATTCGCCAGCTAGGAGCGCCATCAGTAA 8868
 QY 901 TGTTCCTTGGAGTTAAACGAGCTAGAAACAAACCGGAGCTGGGATCAACATATA 960
 |||||||
 Db 8867 TGTTCCTTGGAGTTAAACGAGCTAGAAACAAACCGGAGCTGGGATCAACATATA 8808
 QY 961 ACCGATCAGACATATGCTCCCTGGCTGGAGTAGGTGTACGATTCMACTGCTACTAT 1020
 |||||||
 Db 8807 ACCGATCAGACATATGCTCCCTGGCTGGAGTAGGTGTACGATTCMACTGCTACTAT 8748
 QY 1021 CGTTCAGATAGAGCCCTTATACCTACCTGGGACAAATATCTCTTTATTAACCTGGATAG 1080
 |||||||
 Db 8747 CGTTCAGATAGAGCCCTTATACCTACCTGGGACAAATATCTCTTTATTAACCTGGATAG 8688
 QY 1081 GGGCTACTTGGCAACATCACTTGCATATAGCCCTGCACAAATCGGCAACGTCGCCAGC 1140
 |||||||
 Db 8687 GGGCTACTTGGCAACATCACTTGCATATAGCCCTGCACAAATCGGCAACGTCGCCAGC 8628
 QY 1141 CGTGAACACTCTAAACAAGTGGAGCTCAACCGCGCCAAAGCTTTCCTAAATGGTCTAGG 1200
 |||||||
 Db 8627 CGTGAACACTCTAAACAAGTGGAGCTCAACCGCGCCAAAGCTTTCCTAAATGGTCTAGG 8568
 QY 1201 AATGACTACCCCAAGTATCTACTCAATCCATTTCAAGTAAACACACCGAATCAGA 1260
 |||||||
 Db 8567 AATGACTACCCCAAGTATCTACTCAATCCATTTCAAGTAAACACACCGAATCAGA 8508
 QY 1261 CAAATAATATGAGACAGTATGTAAGATGGCTGCTCTACCTGCTTTCGCAATGG 1320
 |||||||
 Db 8507 CAAATAATATGAGACAGTATGTAAGATGGCTGCTCTACCTGCTTTCGCAATGG 8448
 QY 1321 TGAAGTACTATTAACCAATGATATATCCATTAAGTCTCTTATGATGGAGTGAATA 1380
 |||||||
 Db 8447 TGAAGTACTATTAACCAATGATATATCCATTAAGTCTCTTATGATGGAGTGAATA 8388
 QY 1381 AGAGTTCTATATGTCGAACTCGTCCCATGAGAGAAACGACAGCTATATGATGACCGA 1440
 |||||||
 Db 8387 AGAGTTCTATATGTCGAACTCGTCCCATGAGAGAAACGACAGCTATATGATGACCGA 8328
 QY 1441 CATGATGAAACACTGTTGATATGAACTGAGACGAAATGCCATCTGTTGGCTGCC 1500
 |||||||
 Db 8327 CATGATGAAACACTGTTGATATGAACTGAGACGAAATGCCATCTGTTGGCTGCC 8268
 QY 1501 TCAGGCTGTAAACAGAGAACTCTAATATACAGAGAGGAAATTTGAAAACCATCA 1560
 |||||||
 Db 8267 TCAGGCTGTAAACAGAGAACTCTAATATACAGAGAGGAAATTTGAAAACCATCA 8208
 QY 1561 GACCTCTCATTTGTAGCACTGATGAATATTTGCTGCTATACGCGTAAATATTCAT 1620
 |||||||
 Db 8207 GACCTCTCATTTGTAGCACTGATGAATATTTGCTGCTATACGCGTAAATATTCAT 8148
 QY 1621 GGGCTATGAGCAGGCTATTTACCGCTGACACCACTGTAGGCAATGGCCTTAGCGT 1680
 |||||||
 Db 8147 GGGCTATGAGCAGGCTATTTACCGCTGACACCACTGTAGGCAATGGCCTTAGCGT 8088
 QY 1681 CGCTGCAAAAGTTTACCGCTTATATGATGACCTTACCTGCTGAAGAGCAATTCAGAGA 1740
 |||||||
 Db 8087 CGCTGCAAAAGTTTACCGCTTATATGATGACCTTACCTGCTGAAGAGCAATTCAGAGA 8028
 QY 1741 TTGGAATATACAGAGGCGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTCTCG 1800
 |||||||
 Db 8027 TTGGAATATACAGAGGCGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTCTCG 7968
 QY 1801 TTCTAGCTGGAACCTACCTGCTCCACAACACCCCACTCACTAAAGTTCAACCTCATC 1860
 |||||||
 Db 7967 TTCTAGCTGGAACCTACCTGCTCCACAACACCCCACTCACTAAAGTTCAACCTCATC 7908
 QY 1861 ATCAGATAGTTCACTTACAGTCTAGCTCAACCATCCAGCAACAATAATATAGTACAG 1920
 |||||||
 Db 7907 ATCAGATAGTTCACTTACAGTCTAGCTCAACCATCCAGCAACAATAATATAGTACAG 7848
 QY 1921 TACCAATCTTAACAATAATACGAACAATCAATTAACAACCCCTGATCAACAATAATCAGAA 1980
 |||||||

Db 7847 TACCAATCTTAACAATAATACGAACAATCAACCCCTGATCAACAATAATCAGAA 7788
 |||||||
 QY 1981 TCGTCAACGACACACCA 1999
 |||||||
 Db 7787 TCGTCAACGACACACCA 7769
 |||||||

RESULT 3
 AAT08027
 ID AAT08027 standard; DNA; 2048 BP.
 XX
 AC AAT08027:
 XX
 DT 03-DEC-1996 (first entry)
 XX
 DE S. pneumoniae penicillin binding protein 1A soluble variant DNA.
 XX
 KW penicillin binding protein; PBP 1A; bifunctional protein;
 KW transglycosylase; transpeptidase; identification; assay; inhibitor;
 KW antibiotic resistant; bacteria; soluble variant; protein structure;
 KW X-ray crystallography; determination; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2046
 FT /*tag= a
 XX
 PN GB2290792-A.
 XX
 PD 10-JAN-1996.
 XX
 PF 29-JUN-1995: 95GB-0013306.
 XX
 PR 24-NOV-1994: 94SE-0004072.
 PR 01-JUL-1994: 94IN-0000580.
 XX
 PA (ASTR) ASTRA AB.
 XX
 PI Balganesch TS, Town CM.
 XX
 DR WPI: 1996-042232/05.
 DR P-PSDB: AAW04359.
 XX
 XX Sol. derivs. of bifunctional penicillin binding protein (BPP) -
 PT opt. lack transglycosylase activity, useful to identify and assay
 PT for antibodies or cpds. which bind BPPs
 XX
 PS Claim 11: Pages 65-68; 108pp; English.
 XX
 CC The present sequence encodes a soluble variant of the S. pneumoniae
 CC penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino
 CC acid residues of the wild type protein. Wild type PBP is a
 CC bifunctional protein, which binds the cell membrane when expressed
 CC in a bacterial cell, having transglycosylase and transpeptidase
 CC activities. The variant protein (NCIMB 40665) in conjunction with
 CC a labelled anti-bifunctional PBP monoclonal antibody, can be used
 CC to identify and assay for cpds. which bind bifunctional PBP. Such
 CC cpds., as inhibitors of bifunctional PBP have a potential use in
 CC therapeutic cpds. which inhibit the growth of antibiotic resistant
 CC bacteria. The soluble variant may also be used in X-ray
 CC crystallography.
 XX
 SQ Sequence 2048 BP; 667 A; 498 C; 390 G; 493 T; 0 other;

Query Match 99.1%; Score 1981.4; DB 17; Length 2048;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 1 TAAATCTAGACATATAAATCACTCATTTGCTGACTTGGGTTCTGAGCGCGCTCAA 60
 |||||||

Db 48 taaactctacgcaataaatacatcatctgctgacttggttcttgaagccgcgtca 107
QY 61 TCCCAAGCTATGATTTCCACAGATTTGGTTAAGCAATGCTTTCATGAGACCA 120
Db 108 tgcacaagctaatgatatccacagattggttaaggaaatcglttctatcgaagacca 167
QY 121 TCGCTTCTTCGACCAAGGGGGATTGATACATCCGTATCTGGAGAGCTTTTCGCA 180
Db 168 tgcgtctctcgacccaaggggagttgataccatccgtatccggagagcttcttcgcgaa 227
QY 181 TCTGCAAAAGCAATTCCTTCAGAGTGATCAACTCCACCAAGTTGATTAACTTGAC 240
Db 228 tctgcaaaagcaatccctcccaagtggtatcagcttccatcaacagtlattaaattgac 287
QY 241 TTACTTTTCAATTTGACTTCGACAGACTTTTCTCGTAAGGCTCAGGAAGCTTGGT 300
Db 268 ttaacttcaactcgaactccgaccagactatctctgtaaggtccggaagcttgglt 347
QY 301 AGCGATTGATTAGAAAAGCAACGACGAGAAATCTTGACTACTATATAATTA 360
Db 348 agcgattcagtttagaacaataagcaaccaagaataatcttgactactataataa 407
QY 361 GGTACTACATGCTAATGGGAAGTGAATGCAGACAGCAAGCTCAAAACTACTATGTAA 420
Db 408 ggtctacatgcttaatggaactatggaatgcagacagagctcaaaactactatgtaa 467
QY 421 AGACTTCATAATTTAAGTTTACTGCTAGTTAGCTTGGTGGTGGAAATGCTCAGGCAC 480
Db 468 agactcctaataattaaatttaacccagtlagcttgcgtggtatggtcccaagcac 527
QY 481 AAACCAATATGACCCCTTATTCATTCACATCCAGAACGCCCAAGCCGCAACTTGCTTT 540
Db 528 aaaccaataatgacccctattccatcccgaaagcagcccaagccgcgcgaacttggctt 587
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTAAATAC 600
Db 588 atctgaatagaaaaatcaaggtacatactctgtcgaacagtaatgaaagcagtaaac 647
QY 601 ACCAATTTCTGATGACATCAAAAGTCTCAAAATCAGCAAGTATTTACCGCTTAATGGA 660
Db 648 accaatctatgtaggtctacaaagttccaatccaagaatlaattccctgcttaactgaa 707
QY 661 TAAATACCTCAAGGAAGTCATCATCAAGTTGAAGAAAGAGGCTATTAACCTTCTCAC 720
Db 708 taataactccaaggaagtcataccaagttaagaagaataaagctataaactctacc 767
QY 721 AACTGGATGATGCTCTACACAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
Db 768 aactggatgtagtgtctacaacaatgtagaccaagaagctcaaaaacatctgtggat 827
QY 781 TTACAAATCAGACGAATACGTTGCTTATCCAGACGATGAATTGCAATGCTTACCAT 840
Db 828 ttaacaatacagcaatcgttgcctatccagaatgaatgaatgcaatgcttctcaact 887
QY 841 TGTTCATGTTTCTAAGGTTAAAGTCAATGGCCAGTAGAGGACCCCATCACTCAAGTAA 900
Db 888 tgttgaatgttcttaacggtlaaagtcataatgccagctagagagccatcagtaagtaa 947
QY 901 TGTTCCTTGGGAATTAACCAAGAGTAGAAACAAACCGGCACTGGGATCAATATGAA 960
Db 948 tgttctcttcggaatttaaccaagcagtagaacaacaacgcgcgtgggatacaactatgaa 1007
QY 961 ACCGATTCACAGACTATGCTCCTGCTTGAGACGGTGTCTACGATTCACATGCTACTAT 1020
Db 1008 accgatccaagactatgctcccttgagtagagtgltctcagatcaactgcacactat 1067
QY 1021 CGTTACGATGAGCCCTTAACTACCTCGGACCAATACTGTTTAACTGGGAATAG 1080
Db 1068 cgttaacagatgagccctataactaacctcggtgagcaataaccctgttataacttggatag 1127
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAAATGGGAAAGCTCCAC 1140
Db 1128 gggctactttggcaacatcaacttgcgaataagccctgcgaacaatcggaaggtccagc 1187

QY 1141 CGTGGAAACTTAAACAAAGGTGAGCTCAACCGCGCAAGACTTTCCTAAATGCTTAG 1200
Db 1188 cgtggaaactccaacaagaagtcgagctcaacccgcgcgaagacttctcaaatggtctcg 1247
QY 1201 AATGCACTACCCAAAGTATTCATCTACTCAATGCAATGCAATTCAGTAACACACGGAATGCA 1260
Db 1248 aatgcactaccgaatattcactactcaaatgcatcttcaatgaatlaacacacccgaatcga 1307
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAAGTGGCTGCTGCTTACGCTTTCGCAATG 1320
Db 1308 caaaaaatattgagaagaatgaaagaatggtcgtcgttcaagcttgccttcgcaatg 1367
QY 1321 TGGCACTTACTATTAACCAATGATATCATTAATGATGCTTATGATGAGGAGGAA 1380
Db 1368 tggaaacttactataaaccgaatgatalcatcaataaagtcgtctttagtgaaggatgaa 1427
QY 1381 AGAGTTCTCTAATGTCGGAAGTCTGCCATGAGGAAGACGACGCTTATGATGACGA 1440
Db 1428 agagttctcttaatgctcgagactcgtccatgaaagaagaagacagctataatgtagac 1487
QY 1441 CATGATGAAAACAGTCTTGAATGAACTGACGAAATGCTTATCTTGGCTGCC 1500
Db 1488 catgataaaacagctctgagtagaacctgagcgaatgacctatctgtctgctccc 1547
QY 1501 TCAGGCTGTTAAACAGAGACCTCAACATACAGACGAGGAATTTGAAACACATCA 1560
Db 1548 tcaaggctggtlaaacaagaacctcaactatacagaacggaatctgaaacacacatca 1607
QY 1561 GACCTCTCAATTTGAGCACTGATGAATATTTGCTGCTGCTTACGCTTAAATTAAT 1620
Db 1608 gacctctcaattttagcaacccgtagaacatttgcgtctataagcgttaaatctcaat 1667
QY 1621 GCGTATGAGCAAGGCTATTTAACCGTCTGACACCACTTGTAGCAATGCGCTTACG 1680
Db 1668 ggcgtatgagcaagctatctcaaccgctcgaacacacttgaagcaatggtcctta 1727
QY 1681 CGCTGCCAAAGTTTACCGCTTATGATGACCTACCTGCTGGAAGAACATCCAGAGA 1740
Db 1728 cgcgtccaaagttaacgcgtctatgtagaacctcgtcttgaaggaagaatcccgagaa 1787
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAATAATGTCGTC 1800
Db 1788 ttggatatataccagggggtcttaacagaatggaatctgatttaaaaaatggtgctcg 1847
QY 1801 TTTACGTGGAAGCTCACTGCTCCACAACACCCCATCACTGAAGTTCAAGTCAATC 1860
Db 1848 ttctcagtgagctcaaccgtctccacaacaaccccatcaactgaaagtccaagctatc 1907
QY 1861 ATCAGATAGTTCACATTCACACTGATCAACCACTCCAACGACCAAAATTAATGATAC 1920
Db 1908 atcagatagttcaacttcaacagctcagctcaaccactccaagcacaataatagtagac 1967
QY 1921 TTACCAATCCTTAATTAATACGCAACAAATCAATCAACCCCTGATTCACAATAATACAGA 1980
Db 1968 taccatctcctaacaataatagcaacaatcaaatcaaaccttgatccaacaaatccagaa 2027
QY 1981 TCTCTAACCGACACAACCA 1999
Db 2028 tctctcaaccagcaaca 2046

RESULT 4
AAH02055
ID AAH02055 standard; DNA; 2160 BP.
AC AAH02055;
XX
XX
XX 24-JUL-2001 (first entry)
XX
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2048.
XX
XX Species specific; genus specific; family specific; probe; detection;

KW identification; algal; archaeal; bacterial; fungal; parasitical;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX Streptococcus pneumoniae.
 OS
 PN MO200123604-A2.
 PD 05-APR-2001.
 PF 28-SEP-2000; 2000WO-CA01150.
 PR 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI: 2001-24506/25.
 XX
 PT Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitical species in a test sample -
 PS Disclosure; Page 1463-1464; 1580pp; English.
 XX
 CC The present invention describes a method for generating a repository of
 CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (II)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SQ Sequence 2160 BP; 697 A; 522 C; 413 G; 528 T; 0 other;

Db 279 tccgtcttcgcacacaaagggatgtatccatccgtatccctggaggtcttcctgcgcaa 338
 Oy 181 TCTGCAAGCAATTCCTCCCAAGGATGATCTCACTCCCAAGCTGATTAACTTAC 240
 Db 339 tctgcaaaagcaattcccccaaggtgagtcagctccctcaacaggtgattagttgac 398
 Oy 241 TTACTTTTCAACTTCGACTTCGACAGCTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
 Db 399 ttactttcaacttcgacttcgcagcagacatacttctcgttaagtgctcgaagcttgct 458
 Oy 301 AGCGATTGATTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACTCTATATTAATAA 360
 Db 459 agcgattcagtttagaacacaaagcaacaaagcaaaacttgaccctcatataataaa 518
 Oy 361 GGTCTACATGCTTAATGGGAAGCTATGAAATGACAGACAGCTCAAAACTACTATGTAA 420
 Db 519 ggtctacatgctcttaagtgaactatgaaatgcagacagcagctcaaaactctatgttaa 578
 Oy 421 AGACCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
 Db 579 agacctcaataatttaagttacacttagcttagccttgctgctggaatgctcagcagcc 638
 Oy 481 AAACCAATATGACCCCTTTCACATCCAGAGACGCCCAAGCCGCAAACTTGCTGT 540
 Db 639 aaaccaatataaccctcatlccacatccagagcagcccaagccgcaaaacttgctct 698
 Oy 541 ATCTGAATGAATAATCAAGGCTACATCTGCTGTAACAGATGAGAAAGCAATATAC 600
 Db 699 atcggaaatgaaataacaaaggtcatatcttcgcggaacagtatgagaagaagtcataac 758
 Oy 601 ACCAATTACTGATGACTATCAAAAGTCTCAAAATGACGAAGTAATTAACCTGCTTACATGA 660
 Db 759 accaatctatgctggtcacaagaagtcacaaatgcagaagtaataacctgttaccatgaa 818
 Oy 661 TAATTACTCAAGAGATCATCATCAAGTTGAAGAGAAACAGGCTTAACTACTAC 720
 Db 819 taattacctcaaggaagtcacaaatcaagttgaaagaaacaggtcataactactcac 878
 Oy 721 AACTGGATGATGCTCTCAACAAATGTAAGCAAGAACTCAAAACATCTGTTGGATAT 780
 Db 879 aactggatgatagtctcacaataatgtacgaagaagctcaaaactcctgtggatat 938
 Oy 781 TTACAATACAGACGAATACGTTGCTATCCAGACGATGAATGCAAGTGGCTTACCAT 840
 Db 939 ttacaatacagaagatactgtcctacacagacgatgaattgcaagtcgcttccatcat 998
 Oy 841 TGTGTAGTGTCTCAAGGTAAGTCATGCCCAGCTAGAGACAGCACTCACTCAAGTAA 900
 Db 999 tgtgtatgttctcaggtlaaagtcattgcccagctagagacgccaatcagtcagaagtaa 1058
 Oy 901 TGTTCCTTCGGAATTTAACCAAGCAGTAGAAGAAACCGCAGCTGGGATCACTATGAA 960
 Db 1059 tgttccctcgtgaattacccaagaagtagaagaaacacgcgacccgggatacactatgaa 1118
 Oy 961 ACCGATACAGACTATGCTCTGCTGCTTGAGATACGGTGTCTACGATTCACACTACTAT 1020
 Db 1119 accgatcacagactatgctcctcgttgagtagtgltcagtagtcaactgacactat 1178
 Oy 1021 CGTTCAGATGATGAGCCCTATATACCTCGGAGCAAAATACCTGTTATATCGGAGTAG 1080
 Db 1179 cgttcagatgagccctataactacccctggaacaaataacccgttctataactgtagag 1238
 Oy 1081 GGGCTACTTTGGCAACATCACTTGCAATAGCCCTGCAACAACTGCGAAGCTGCCAGC 1140
 Db 1239 gggctactttggcaacatcaacttgcataacgcccgtgaaacaaatcgggaaaacgtcccaagc 1298
 Oy 1141 CGTGAATACTCTAAACAAGGTGAGACTCAACCGGCGCAAGCTTCTTAATGGTCTAG 1200
 Db 1299 cgtggaactcttaacaagtcgagacccaacgcgcaagccttcttaaatgtgtcgg 1358
 Oy 1201 AATGACTACCCCAAGTATCTACTACTCAATAGCAATTTCAAGTAAACCAACCGAATGAGA 1260
 Db 1359 aatgactaccaccaagtatctactactcaaatgccaatttcaagtaacacaacacgaatcaga 1418

```
QY 1261 CAAAAATATGAGACAGTAGTGAAGAAGTGGCTGCTGCTTACCGCTGCTTGGCAATGG 1320
Db 1419 caaaaalatalgagcaagtagtgaaagatggtcgtctacgcgtccttgcgaatg 1478
QY 1321 TGAACCTTACTATTAACCAATGTTATATCAATAAAGTGTCTTATGATGAGATGAAAA 1380
Db 1479 tgggaactactataacaatagtatcatcaataaagtcgtctttagtgatggtgagtgaaaa 1538
QY 1381 AGACTTCTTAATGTCGGAACCTGTCGATGAGAAAAAGCAAGCATATATGATGACCGA 1440
Db 1539 agagttctcctaattgtaggaactcgtgcataagaaagcagcagctatatgatgacgga 1598
QY 1441 CATATATAAAAACAGCTTCTGACTTATGAGACGAGCAAAATCCCATATCTGTTGGCTGCC 1500
Db 1599 calgataaaaaacagctcttgatgatalgagaatgagaatctgctctggtcc 1658
QY 1501 TCAGGCTGTGTAACAGGAACCTTACTATATACAGAGGAAATGTAACCAATCAATCA 1560
Db 1659 tcaggtcgtgtaaaaaacagaaaccttaactatacagacgaggaattgaaaaaccatcaa 1718
QY 1561 GACCTCTCAATTTGTAGACCTGATGACTATTTGCTGGCTATACGGCTAAATATTCAT 1620
Db 1719 gacctccaatttgtagcacctgtatgaactattgtcgtcctatacgctaaatataatcaat 1778
QY 1621 GGGCTGTATGACAGCGCTATTTCTAACCGTCTGACACCACTTGTAGCAATGGCTTACGGT 1680
Db 1779 ggcgtgatgagacaggtcattcttaacgcgtctgacacccactttaggcaatggtcctacggt 1838
QY 1681 CGGTGCCAAGTTTACCGCTCTATGATGACTTACCTGCTGTAAGAGCAATCCAGAGA 1740
Db 1839 cgtgcgtcaaaagtttacccgtctctatgatagactaaccttctgaaaggaagcaatccagaga 1898
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTGTCTG 1800
Db 1899 ttggaatataccagaggggctctacagaaatggaatctgtatttaaaatggtgtcgt 1958
QY 1801 TTTACTGTGACACACCTGCTGTCCACAACAACCCCATCAACTCAATGTAAGTTTACCTATC 1860
Db 1959 ttctaacgtgagtcacacctctccacaacaccccccaactcgaactgaaagttcaaacctc 2018
QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAACGCAACAAATATATAGTACG 1920
Db 2019 atcagatagttcaacttcacagctctagctcaaccatcccaagcaataatagtgac 2078
QY 1921 TACCAATCTTAACAAATATAGCAACAATCAAAATACAAACCCCTGATCAACAAAATCAGAA 1980
Db 2079 taccaatcctaacaataatagcaacaatcaatacaacccctgatacaacaataatcagaa 2138
QY 1981 TCTCAACGACGACACCA 1999
Db 2139 tctcacaaccagcaaacca 2157

RESULT 5
AAS55688 standard; DNA: 2160 BP.
XX AAS55688:
XX AAS55688:
XX 13-FEB-2002 (first entry)
DE Streptococcus pneumoniae DNA for cellular proliferation protein #259.
XX
XX Antisense: ds; prokaryotic cellular proliferation gene;
KM antidiotic; antibacterial; drug design.
XX
XX Streptococcus pneumoniae.
XX
XX MO200170955-A2.
XX
XX 27-SEP-2001.
XX
```

```
PF 21-MAR-2001; 2001WC-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Heselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR P-PSDB; AAU37829.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Claim 27; Seq ID No 9325; 511p; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2160 BP; 697 A; 522 C; 413 G; 528 T; 0 other;

Query Match 99.1%; Score 1981.4; DB 23; Length 2160;
Best Local Similarity 99.4%; Pred. NO. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATATAAATCACTGATGCTGAGCTGGGTTCTGAACGCCGCTCAA 60
Db 159 taaatctacgacaataataataatcaatcattgctgacttgggtctgaaagccgcgtcaa 218
QY 61 TGCCCAAGCTATATATTTCCACAGATTTGGTTAAGCAATGCTTTTATCGAAGACA 120
Db 219 tgccttcgacacaggggattgataccatccgataccctgatacctcgtgagcttcttcgcgaa 278
QY 121 TCGCTTCTTGACACAGGGGATTGATACCATCGTATCCTGGAGGTTCTTGGCGAA 180
Db 279 tcgcttctcgacacaggggattgataccatccgataccctgagagcttcttcgcgaa 338
QY 181 TCTGCAAAAGCAATTCCTCCAAAGGTGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 240
Db 339 tctgcaaaagcaattccctccaaggttgatcagctctcaaccatacagttgataagttgac 398
QY 241 TTACTTTTCAACTTCGACTTCGACAGAGACTATTTTCGTAAGGCTAGAGAGCTTGTT 300
Db 399 ttacttttcaacttcgacttcgacagactattctcgtlaaggtctgaagaaagcttggct 458
QY 301 AGGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATTAATAATA 360
Db 459 aggatcagttagaaacaaaagcaacccaagcaagaaatcttgacctactatataataaa 518
```

QY 361 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGCAAGCTCAAAACTATATGCTAA 420
|||||
Db 519 ggtctacatgcttaatgggaactatggaatgcagacagcctcaaaactactatgtgtaa 578
QY 421 AGACCTCAATAATTTAGTTTAACTCAAGTTAGCTTGTGGCTGGAAATGGCTCAGGCAACC 480
|||||
Db 579 agacctcaataattttagttttaaactcaagcttagccttgctggccttggaatgctcctcagacc 638
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAGACAGCCCAAGACCGCGAAACTTGGCTTT 540
|||||
Db 639 aaaccaatatagacccctattccacatccagaagaccccaagaccgccaacttggtctt 698
QY 541 ATCTGAAATGAAAAATCAGGCTACATCTCTGCTGAACAGTAGAGAAGCAAGCTCAATAC 600
|||||
Db 699 atctgaaatgaaaaatcaagctcacatctctgcgcgaagctatagaagaagcaatcac 758
QY 601 ACCAATATACGATGAGTACAAAAGTCTCAATATGCAATGCAAGTAATTAACCTGCTTACATGA 660
|||||
Db 759 accaatatacgtatggctacaaagctcacaatcagcaagtaataccttcttaccatgga 818
QY 661 TAATTACCTCAGGAAGTCAATCAATCAAGTTGAAGAAGAAAGAGCTATTAACCTACTCAC 720
|||||
Db 819 taattacctcaaggaagctcacaatcagaatggaagaagaacggctataactactcac 878
QY 721 AACTGGAGTGAATGTCTACACAAATGTGACCAAGAAAGCTCAAAAACATCTGTGGATAT 780
|||||
Db 879 aactggatgtagtctcacacaatgtagaccaaagctcaaaaaacatctgtggat 938
QY 781 TTTCAAATACGAGAGAAATACGTTGCTTTCAGACGATGAATTTGCAATGCCCTTCTACAT 840
|||||
Db 939 ttccaataacgaagaaatagctccttaccagacgaatggaatgccttcttaccat 998
QY 841 TGTGTGATGTTTCTAACGGTAAAGTATTTGCCAGCTAGAGCAGCCATAGTCAAGTAA 900
|||||
Db 999 tgtgtatggttcttaacggtaagctatgtccagctagagcgacgccaatgaagctaa 1058
QY 901 TGTTCCTTTCGGAATTTAACAGCAGTAGAACAACCGCGACTGGGATCAATATGAA 960
|||||
Db 1059 tgttccttcggaatttaaccaagcagtagaacaacacgcgactggtggaatcaatagaa 1118
QY 961 ACGGATACGAGACTATGCTCTGCTTGGAGTACGGTGTACGATTCACATGCTACTAT 1020
|||||
Db 1119 accgaatacgaactatgctcctgcttgagctagctgctacagtaacactccatlat 1178
QY 1021 CGTTCAAGATGAGCCCTATTAATACCTGGAGCAAAATACTCTGTTTATTAATCGGATAG 1080
|||||
Db 1179 cgttcaagatgagccctataactaccctgggacaataaccctgtttatacttggatag 1238
QY 1081 GGGCTACTTTGGCAACATCACTTGGCAATAGGCCCTTGCAACAAATCGGAAACGTCCGAG 1140
|||||
Db 1239 gggctactttggcaacatcacctctgcaatagccctgcaacaatcggaagctccagc 1298
QY 1141 CGTGGAAACTCTAAACAAGTGGGACTCAACCGCGCAGACTTTCTTAATAGTGTAGG 1200
|||||
Db 1299 cgtggaaactctaaacaagctggagctcaacgcgcgaagacttctctaaatgctcgcg 1358
QY 1201 AATCGACTACCCAAGTATTCACACTCAATGCGCATTTCAAGTAAACAACCAATCAGA 1260
|||||
Db 1359 aatcgactacccaagatcttaactactcaaatgctcaatlttcaagtaacacaacgaaatcaga 1418
QY 1261 CAAAAAATATGAGCAAGTAGTAAGATGCTGCTCTTAACGCTGCTTTGCAATAGG 1320
|||||
Db 1419 caaaaaatatgagcaagtagtaagaagatgctgtcttaccgtcgtcccttgcaaatg 1478
QY 1321 TGGAACTTACTATAAACCATATATTCATTAAGTCGTCTTAAGTAGTGAGGTGAAAA 1380
|||||
Db 1479 tggaaacttactataaaccatatgatatccataaagctgtctttagtgaatggtgaaaa 1538
QY 1381 AGAGTCTCTAATGTGGAAGCTGTGCCATGAGGAAGACAGCAAGCTATTTGATGACCGA 1440
|||||
Db 1539 agagttctcttaatgtcggaaactcgtgcatagaaagaagacagcctatatatgtgacgga 1598

QY 1441 CATGATGAAAAACGCTTTCATTAATGGAAGTGCAGCAAAATGCTATTCCTGCTGCC 1500
|||||
Db 1599 catgatgaaaaacgcttcttgatgtatggaactgagcgaatgcttatcttcttgctcc 1658
QY 1501 TCGAGCTGTAAAAACAGAACCTCTACTATATACAGACAGGAATAATGAAAAACCATCAA 1560
|||||
Db 1659 tcagcgtgtataaacaagaacctctaacctaacagacgaggaatctgaaaaaccaacaa 1718
QY 1561 GACCTCTCAATTTGTACACCTGTATGAACTATTTGTGCTATTAACGGCTAAATATTCAT 1620
|||||
Db 1719 gactctcaatttgtagcacttgatgactatgactatctgctgctatcaggtataatlatcaat 1778
QY 1621 GCGTGTATGACAGAGCTATTCTTAACCGTCTGACACCACTTGTAGCAATGGCCTTACGGT 1680
|||||
Db 1779 gctgtatgtagcagcgtattcttaacgcttgacacccacttgtagcaatgcttaccgt 1838
QY 1681 CGCTGCCAAGTTTACCGCTATGATGAGACTACCTGCTGGAAGGAAGCAATCCAGAGA 1740
|||||
Db 1839 cgtgcacaagtttaacgcctctatgacttaccctgtcgaaggaagcaatccagaaga 1898
QY 1741 TTGGAATATACAGAGGCGCTCTACAGAAATGGAATTCGTATTTAAAATGCTGCTCG 1800
|||||
Db 1899 ttggaatatcccgaggggctctacagaaatggagaatctgtatttaaaaaatgctcgt 1958
QY 1801 TTCTACGTGGAAGTCACTCTCCACACAACACCCCATCACTGAAGTTCAAGCTCATC 1860
|||||
Db 1959 ttctacgttggagcttaccctgtctccacaacacccccaatcaactgaaagtccaagctcatc 2018
QY 1861 ATCAGATAGTTCAATTTCAAGTCTAGCTGACACACCTCCAGCAAAATATATGTCAGC 1920
|||||
Db 2019 atcagaatagttcaacttccagcttagctcaaacctccagaacaataatagctacgac 2078
QY 1921 TACCAATCCCTAACAAATATATGCGCAACATCAATATACACCCCTGATCAACAATCAGAA 1980
|||||
Db 2079 taccaatcccttaacaataataatgcacaacaatcaaatcaacccctgatacaaaaaatcaaga 2138
QY 1981 TCCTCAACGACACAAACA 1999
|||||
Db 2139 tcctcaacgacacaacca 2157

RESULT 6
AAH02056
ID AAH02056 standard; DNA; 2160 BP.
XX
AC AAH02056;
XX
DT 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2049.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitical;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
XX
PR 19-MAY-2000; 2000CA-2307010.
XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
PI Berleron MG, Bolssinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;

XX WPI: 2001-245006/25.

XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitical species in a test sample -
PS Disclosure; Page 1464-1465; 1580pp; English.

XX The present invention describes a method for generating a repertory of
CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitical
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific, bacterial, fungal
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitical species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.

XX Sequence 2160 BP; 696 A; 495 C; 402 G; 567 T; 0 other;

Query Match 83.9%; Score 1677.4; DB 22; Length 2160;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 1798; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATATAATACATCATTTGGTTCGTAAGCGCGGCA 60
DB 159 taaatctacgacataataatacaatcactgctgactggttcgaaacgcgcgca 218
QY 61 TGCCAGCTAATGATATTCACAGATTGTTAAGCAATCGTTCTATCGAGACA 120
DB 219 tgcacaagctaatgatatctccacagattggttaagcaatcgcttctatcgaaagca 278
QY 121 TCGCTTCTTGACACAGGGGGATTGATACATCCGATCTTGAGCTTTTGGGCAA 180
DB 279 tcgcttcttgacacaggggattgattccacgcglatcttgaggcttcttcgcgca 338
QY 181 TCGCAAGACATTCCTCCCAAGTGGATCACTCCACCAAGTGGATTAAAGTTGAC 240
DB 339 tcgcaaaagttaattccctcccaagggtgatactccaccaacagtttgtaagtgcac 398
QY 241 TTAATTTCAACTTGACTTCGACAGACTATTTCTGTAAGCTTCAGGAAGCTGGTT 300
DB 399 ttactttcaactcgacttcgaccagactatttctcgtaaggtctcagaagcttggtt 458
QY 301 AGCGATTTCAGTTGAAACAAAACCAAGCAAGAAATTTTACCTACTATTTAAATTA 360
DB 459 agcgatttcagttgaaacaaaacaaacaaacagaaatcttgacctactataataaa 518
QY 361 GGTCTACATGTCTAATGGAGATGGAATGAGAGAGAGAGCTCAAAATACTATGATGTA 420
DB 519 ggtctacatgtctaatggagcactatggaatgcagacagcgtccaaaactactatgttaa 578
QY 421 AGACCTCAATTAATTAATTTACCTCAGTTAGCTTGCTGGTGAATGGCTCAGGCACC 480
DB 579 agacctcaatatttaagttaacctcagttagcttgctgagatgtcctcagagacc 638

QY 481 AAACAATATGACCCCTATTACATCCAGAGACGCCCAAGACCGGAAACTGTGCTT 540
DB 639 aaacaaatatagacctattcaatccagagaagcgcgaagccggaactgtgtctt 698
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAAGCACTATGAGAAGCACTCAATAC 600
DB 699 atctgaatgaaaaaatcaaggttacatctctgtcgaacagtatgagaagcagttcaatc 758
QY 601 ACCAATTACTGATGAGCTACAAAGTCTCAAAATACGACAGTAATTTACCTGTCATGCA 660
DB 759 accaattactgagctacaaagtcctcaaatcagcaagtaatacccttgcttaacatgga 818
QY 661 TAATTACCTCAGAGAGATGATCAATCAAGTTGAAGAAACAGGCTATACCTACTAC 720
DB 819 taattacctcaagaaggtctacatccatcagaagtagaacaagaacggtcttaaccttaac 878
QY 721 AACTGGATGATGCTATACCAAAATGTAGACCAAGAAAGCTCAAAATCATCTGTGGATTA 780
DB 879 tactggatgagtgatttaacaaatgtagaccagaagcctcaaaaactatctggtgat 938
QY 781 TTTCATATCAGACGAATAGCTTGCTATCCAGACGATGAATTCGACAGTCCCTTACCAT 840
DB 939 ctacaactcagatcaatagctctctaccctgacagatgttgcaagtcgacatcaggt 998
QY 841 TGTGTAGTTTCTTAACGTTAAGTCAATTCGCGAGCTAGAGCAAGCCATCATGTCAGTAA 900
DB 999 cgtagagtttcttaacggttaagttcaatgcgccaacttgagagctcgttacaagaagtaa 1058
QY 901 TGTTTCTCTGGAATTAACCAACAGTGAAGAAACAAACCGGAGCTGGGATCAATTAAGA 960
DB 1059 cgtttctcttgtagcaacaaagcttgtagaacaacatcgagactggttctctgtagaa 1118
QY 961 ACCGATCAGACACTATGCTCTGCTGTGAGTAGTACGTTCTACGATCAACTGCTACTAT 1020
DB 1119 accgatcagcatatgacactgacactgacatagatacagtggttaatgattccactgcaactat 1178
QY 1021 CGTTACGATGAGACCCCTATATACCTCTGGGACAAATATCTCTGTTATATACCTGGATAG 1080
DB 1179 ggttaatgataatctcttaatactatccggaagaagacacgttcttaacactggtgagaa 1238
QY 1081 GGGCTACTTTGGCAACATCACTGTCATACGCGCTTCACAAATCCGGAACGCTCCAGC 1140
DB 1239 agcatatttggtaataataatctctgcaatagatgcttcttaacaatacgaatagtcaagc 1298
QY 1141 CGTGAATCTTAACCAAGGTGAGCTCAACCGCGCCAGACTTTCTTAATAGGTCTAGG 1200
DB 1299 cgttgagacttgataaagtcgtctagatagagctaaacacttcttaagtgtctg 1358
QY 1201 AATCGACTACCAAGTATTCATCACTCAATATGCCATTTCAGTAACCAACCAATACGA 1260
DB 1359 tctgactatcccaagcattgataaagaagccattcaagtaataacaaagataaa 1418
QY 1261 CAAAAAATATGAGCAAGTATGTAAGGATGAGTGGCTGTAGCGCTCTTGGCAATGG 1320
DB 1419 taaacaatacagggcaagtagtgaaaaaatggtcgtcgtcttaagtccttgcaaatg 1478
QY 1321 TGAACCTTACTATTAACCAATGTATATCATTAAGTGTCTTATAGTATGAGTGAATA 1380
DB 1479 tggcaactactataaacaatgatatcaataaagtcgtctcagtgatgaagtaaaaa 1538
QY 1381 AGAGTTCTCTAATATGTGGAAGTCTGTCATGAAAGGAAGCAACAGCCATATATGATGACGA 1440
DB 1539 agagttctcttaatgtcgaagctcgttgcaaggaagacagcctataatgtagcga 1598
QY 1441 CATGATGAAAAAGTCTTGAATGATGAGAACTGGAACGCAATATCTTGTGGCTCC 1500
DB 1599 catgatgaaaaagctcttgactatggaactgggcgtgagagcctatcttcttggtctcc 1658
QY 1501 TCAAGCTGTGTAACAGGAAGCTCTTAATCTATACAGACGAGAAATGGAACCAATCAAA 1560
DB 1659 tcaagctgtgtaaaacaggaaccttaactatacagatgagaagatgtgaaaaacacatcaa 1718

QY	1561	GACCTTCACATTTTGTAGCACCTTATGACATATTGCTGGCTATACGGCTAAATATTCAAT	1620
Db	1719	gaacacgycgataibtagctccagaibgaaiygtitgtgltactcgtlaaglatctcat	1778
QY	1621	GGCTGTATGGACAGCCTATTTCTAACGGCTCGACACACCTTGAAGCATTGGCGCTTAAGCGT	1680
Db	1779	gycgtatibgacacggtatctcgatctglttaactccctacgtcttgagatbgtltccagtt	1838
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGTATGACCTACCTGTCTGAAGAGCAATTCAGAGA	1740
Db	1839	tgcgcgtcaaaagtattatcgctccaaibgaibtaacglatctcatcaagaaga tactcatccagaga	1898
QY	1741	TTGGAATATACAGAGGGGGCTCTACGAAATGGAGATTCGTATTTAATAATGGTGGCTCG	1800
Db	1899	cttgcagatgcibccagaaacgaccttccagaaibggagatctglatttaaaatbgtgtccg	1958
QY	1801	TTTACCTGGAGACCTACCTGCTCTCCACACACACCCCATCACTGAATTAATTACCTCATC	1860
Db	1959	ttctacgttggaacttaccctctccacaacaccccatcaactgaagtltcaagctcatc	2018
QY	1861	ATCAGATAGTTCAACTTCACAGTCTTAGCTCAACACCACTCCAGACCAATAATTAAGTACGAC	1920
Db	2019	atcagatagttcaacttccacagctctcgtcccaacacactccaagcacaanaibtaibtagtac	2078
QY	1921	TACCAATCTCAACAAATTAATACGCACAAATCAAAATCAACCCCTGTATCAACAAATTCGAA	1980
Db	2079	tacccaatccctacaacaataatcaggaacaaatccaataatacaacccctgtatcaacaaatctgaa	2138
QY	1981	TCCTCAACACGACACACCA 1999	
Db	2139	tcctcaacacgacacaaca 2157	

RESULT 7

XX	AAH01178 standard; DNA; 2160 BP.
XX	
AC	AAH01178;
XX	
DT	24-JUL-2001 (first entry)
XX	
DE	Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1169.
KW	Species specific; genus specific; family specific; probe; detection;
KW	identification; algal; archaeal; bacterial; fungal; parasitic;
KW	microorganism; diagnosis; translation elongation factor Tu; toxin;
KW	translation elongation factor G; RecA recombinase; resistance;
KW	catalytic subunit of proton-translocating ATPase; antimicrobial;
KW	vaccine; primer; ds.
XX	
OS	Streptococcus pneumoniae.
PN	WO200123604-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000MO-CA01150.
XX	
XX	28-SEP-1999; 99CA-2283458.
PR	19-MAY-2000; 2000CA-2307010.
XX	
PA	(INFECTIO DIAGNOSTIC (IDI) INC.
XX	
PI	Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
XX	Picard FJ, Roy PH;
XX	
DR	WPI; 2001-245006/25.
XX	
SS	Nucleic acid sequences are used to generate universal probes and PT primers which can be used to identify and detect the presence of algal PT algae, bacterial, fungal and parasitcal species in a test sample - XX disclosure; Page 1044-1045: 1580bp; English.

The present invention describes a method for generating a repertoire of nucleic acids of *tuf*, *fus*, *atpB* and/or *recA* genes from which probes and/or primers are derived. The method comprises amplifying the nucleic acids of determined algal, archaeal, bacterial, fungal and parasitical species with a combination of defined primer pairs. The method can be used for producing probes and/or primers for detecting one or more related microorganisms e.g. algae, archaea, bacteria, fungi and parasites, for universal detection and for specific and ubiquitous detection and identification of an algal, archaeal, bacterial, fungal and parasitical species, genus, family and group. A nucleic acid (1) obtained using the method of the invention can be used for the universal detection of any bacterium, fungus or parasite in a sample and for the detection of at least one antimicrobial agent resistance gene or at least one toxin gene. hexa nucleic acids are used for the specific and ubiquitous detection and for identification of *Streptococcus pneumoniae*. (1) can be used to design a therapeutic agent which is effective against microorganisms. Microbial species or genus or family or phylum or group which can be detected include *Albortrophia adiacens*, *Bordetella sp.*, *Corynebacterium sp.*, *Enterobacteriaceae* group, *Escherichia coli*, *Mycobacteriaceae* family, *Pseudomonads* group, *Streptococcus sp.*, *Neisseria gonorrhoeae* and *Staphylococcus sp.* Using DNA based tests provides faster results than substrate specificity tests as results can be determined in an hour and improved accuracy is also achieved. AAH00010 to AAH002304 represent nucleotide sequences and primers/probes which are given in the exemplification of the present invention.

Query Match	82.2%;	Score 1642.2;	DB 22;	Length 2160;
Best Local Similarity	88.8%;	Pred. No. 0;		
Matches 1776;	Conservative	0;	Mismatches 223;	Indels 0;
			Gaps	0;

OY	1	TAAATCTACGACATATAAAATCAACTCTATTGCTACTTGGGTTCTGAAGCGCCGCTCA	60
Db	159	taaaatctacgcacaaataaaaatacaactcatctgctgacttgggtcttcgaaagccgcgtcaa	218
OY	61	TGCCAAGCTATGTGATATTCCACAGATTGGTTAAAGCAATCGTTTCATATGAAGACCA	120
Db	219	tgccaaagctaatgtatatctccacagatttgglttaagaaacgtttctcatcgaagaacca	278
OY	121	TGCGTTCTTTCGACACACAGGGGGGATTTGATACCTTCGTATCCCTGGAGACTTTCTTGCGCAA	180
Db	279	tcgctcttcgcgcacacagggggaattgatccaatccgtatccgtatccgtgagagctctctgcgcaa	338
OY	181	TCTGCAAAAGCAATTCCTCTCCACAGGTGGATCAACTCTCACCAACAGTTGATTAAGTTGAC	240
Db	339	tctgcaaaagtaattccctcccaagtgatgatacaactctcacccaacagttgatlaaattgac	398
OY	241	TTACGTTTTCACCTTGACGCTTCGACACAGCTATTTCGTAAAGGCTCAGGAAGCTTGGT	300
Db	369	ttaactctcaacctgcgactccgcaccagactatcttcgtlaaggtccaggaagcttggct	458
OY	301	AGCGATTACAGTTAGTAACAAAAAGCAACACAGCAAGAAATCTTGACCTACTATATAATTA	360
Db	459	agcgatttcagtttagaacaacaaaagcaaccaacaaagaaatcttgactactatataataaa	518
OY	361	GGTTCATGTCCTAATGGGAACTATGGAATGACAGACAGCTCAAAATCTACTATGTAA	420
Db	519	ggtctacatgtctaatagtcgaactatgaaatgaaagtcagcagctcaaaactactatgtaa	578
OY	421	AGACTCATAATTTTAAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGACAC	480
Db	579	agactcataaattaaagtttaacctcagtttagcctctgcttggtatgtaagtcctcaggcacc	638
OY	481	AAACCAATATGACCCCTATTCACATCCGAAAGACACCCCAAGACCCGCAAAATTTGCTTT	540
Db	639	aaaccaaatatgacccctatccaatccaagaacagcccaagccgcgcgaacttggctt	698
OY	541	ATTCGAAATGAAAAATCAAGCTACATCTCTGCTCAACAGTATGAAAGCAAGTAAATAC	600
Db	699	atctgaaatgaaaaatcgaaggtatatactctcgtctaaacagatgaaaaagcgtataaattac	758

QY	601	ACCAATTACGATGAGCTACCAAGTCTCCAAATACGACAGTATTTACCTGCTTACATGGA	660
Db	759	accgaattacgtagtgagctacaagaagcttcaactcagaagtaatactacccgtcttaca1gga	818
QY	661	TAAATTACCTCAAGGAAGTCTCATCAATCAATTGACAGAAACAGAGCTATTAACTACTAC	720
Db	819	taattactactaaggaggtcactcaactaagtagaacaagaagaactgctataataccttcaac	878
QY	721	AACCTGGATGGATGTCTCACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
Db	879	tactcggagtagagtagtctacacaatgtagaaccagaagctcaaaaaca1ctg1ggagat	938
QY	781	TTACAAATACAGAGGAATACGTTGCCATATCCAGACGATGTAATTGGCAAGTGGTTCACCAT	840
Db	939	ctacaactccga1ccaataagcttcttaaccctgagagatggtttgcaag1cgcataccaagcgc	998
QY	841	TGTTGATGTTTCTTAACGGTAAAGTCAATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA	900
Db	999	cgtagagtgcttccaat1ggttaagtagtctgcgccact1ggag1ctg1tcaccaaga1cagtaa	1058
QY	901	TGTTTCTCTGGGAATTAACCAACCACTATAAACAACCGCAGCTGGGATCACTATGAA	960
Db	1059	cgcttca1cttgg1tacaacaacaagc1g1tggaaacaacac1g1gag1c1g1g1ctg1ctatgaa	1118
QY	961	ACCGATACAGAGATATGCTCTCTGCTTGGAGTAGACGGTGTCTACGATTAACGTGTACTAT	1020
Db	1119	accaatcaccga1tat1gcaac1g1cga1taaga1tag1gt1ttag1tgc1tccact1gcaac1tat	1178
QY	1021	CGTTTACAGATGAGCCCTATATACTACCTCTGGGACAAATACTCCTGTTTATTACTGGGATAG	1080
Db	1179	ggttaatg1at1cttact1ataact1cc1gg1aa1caag1caac1c1gt1ct1caact1gg1atag	1238
QY	1081	GGGCTACTTGGCAACATCACTTCCATGGAATATAGCCCTGGCAACATGCGAAAGTCCCAAGC	1140
Db	1239	agc1at1t1c1g1t1aa1tat1ct1aact1t1c1g1t1aa1t1g1c1t1t1caac1t1ca1ga1at1g1t1caac1g1c	1298
QY	1141	CGTGGAACTCTTAACAAGCTCGAGCTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG	1200
Db	1299	cg1t1ga1g1act1t1ga1ta1a1g1t1c1g1t1c1ag1a1g1ta1a1a1c1t1t1ca1t1g1t1c1t1g	1358
QY	1201	AATGCTATACCCCAAGTATTCATAC1CAAC1AATGCGANTTCACATTAACACACCGAATCAGA	1260
Db	1359	1at1c1ga1c1t1a1c1t1aa1c1a1g1t1a1t1a1g1caa1c1g1c1a1t1t1a1a1g1ta1a1a1c1a1a1g1t1c1aa	1418
QY	1261	CAAAAAATATGAGGACAGTAGTGAAGAAAGATGGCTGCTTACGCTGGCTTTGGAAATAG	1320
Db	1419	ta1a1c1a1t1a1c1g1a1g1c1a1g1t1a1g1aa1a1a1t1g1c1g1c1g1t1a1t1g1t1c1t1t1g1a1a1g1g	1478
QY	1321	TGGAACCTACTATTAACCAATGATATTCATATAAGTCGTTTACTGATGGGAGTGAATA	1380
Db	1479	1t1g1ga1c1t1a1c1t1a1a1c1a1a1g1t1a1t1c1c1a1a1a1g1t1c1t1c1a1g1t1a1g1a1a1a1a	1538
QY	1381	AGACTTCTCAATATGTGGGAACCTGTGSCCATGGAAGAAAGACAGCTTATATGATGACCGA	1440
Db	1539	ag1a1g1t1c1t1c1aa1g1t1c1g1a1c1t1c1g1c1a1t1g1c1a1t1g1a1g1a1a1c1c1a1t1a1a1g1c1g1a	1598
QY	1441	CATGATGAAAACAGTCTTGACTTATTTGGAACCTGGAGAAATGGCTATTCTGGTTGGCTCC	1500
Db	1599	ca1t1a1t1g1aa1a1c1a1g1t1c1t1g1a1c1t1a1t1g1a1c1t1g1g1c1g1t1a1c1t1c1t1c1t1c1c1t1c	1658
QY	1501	TCAGGCTGGTAAACAGGAACCTTAAC1ATACAGACGAGGAATTTGAAAAACACATCTAA	1560
Db	1659	1t1c1a1g1c1t1g1t1a1a1a1c1a1g1a1c1c1t1a1c1a1c1a1t1c1a1g1t1g1a1g1a1g1t1g1a1a1c1c1a1t1a	1718
QY	1561	GACCTTCATATTTGTAGCACACTGATGAACATATTTTGTGGCTTATAGCCTTAATATCTAAT	1620
Db	1719	g1a1c1a1c1g1c1t1a1g1a1g1t1a1g1t1c1c1c1a1g1t1c1t1a1c1c1t1c1t1a1c1t1c1t1c1t1c1t1c1t	1778
QY	1621	GGCTGTATGGACAGGCTATTCTTAACCGTCTGACACCAGTGTGAGCAATGGCTTACGGT	1680
Db	1779	g1c1c1t1a1t1g1a1c1a1g1t1a1t1c1g1a1t1c1t1a1c1c1t1c1t1a1c1c1t1g1a1g1t1t1c1t1c1g1t1	1838

QY	1661	CGCGCCAAAGTTTAAOCGCGCTGATGATMGACCTACACCGTCGTGAAGGAAGCAATCCGGAAGA	1740
Db	1839	ltycgccctaaagttlalttcgtctccaatgataacgtlactlctatagaagaatctcatcccgaga	1899
QY	1741	TTTGAAATATACCAAGAGGGGCTGTACAGAAATGAGAAATTGCTATTTAATAAATAGTGCTCG	1800
Db	1889	cttgaacgatgcggaacggaacttttcagaaacggygaatttgcattccaataatggaagctcg	1956
QY	1801	TTTACACTGGCACTCACTCCTCCACACAGACCCCCATCACTGAAGTTTCAGCTCATC	1860
Db	1959	cccaatatgactgaagaaacccctactctcaacaatcctcaacagctgaagtcaagtcac	2018
QY	1861	ATCAGATAGTTCAACTTCACACTGCTAGCTCAACCACTCCAGCACCAATATATAGTACAC	1920
Db	2019	atcagatatgttcaactcttcaacagcttagctccaacactccaagacacaaataatagtagac	2078
QY	1921	TACCAATCCCTAACCAATATATACGCAACCAATCAATATCAACCCCTGATTCAMAAATACAA	1980
Db	2079	taccaatccctaaacataatacgcacaacatcaaatacaacccctgtatcaacaataatcagaa	2138
QY	1981	TGCTCAACGACGACAACCA 1999	
Db	2139	tcctcaaacgcagcacaccca 2157	

RESULT 8

ID AAH02054 standard; DNA; 2160 BP.

AC AAH02054;

DT 24-JUL-2001 (first entry)

DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2047.

KW Species specific; genus specific; family specific; probe; detection;

KW microorganism; diagnosis; translation elongation factor Tu; toxin;

KW catalytic subunit of proton-translocating ATPase; antimicrobial;

XX

XX

XX

XX

XX

PR 19-MAY-2000; 2000CA-2307010.

PA (INF-) INFECTIO DIAGNOSTIC (IDI) INC

PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M,

XX

XX

PT primers which can be used to identify and detect the presence of algal

XX

XX

CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes

CC acids of determined algal, archaean, bacterial, fungal and parasititcal

CC used for producing probes and/or primers for detecting one or more

CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archeal, bacterial, fungal
CC and parasitic species, genus, family and group. A nucleic acid (1)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
CC
CC
CC
SQ Sequence 2160 BP; 701 A; 494 C; 395 G; 570 T; 0 other;

Query Match 81.1%; Score 1621.4; DB 22; Length 2160;
Best Local Similarity 88.2%; Pred. No. 0;
Matches 1763; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATAAATCACTCATGTGCTGACTGGTTCGAGCGCGCTCAA 60
DB 159 taaatctacgacataaaaaatacaactcatctgacttgggttctgaaagccgctcaa 218
QY 61 TSCCCAGCTAATGATATTCACAGATTTGGTTAAGGCAATCGTTTATGGAAGACCA 120
DB 219 tgcacagataatgatattccacagatttggttaagcaatcgtttctatcgaagacca 278
QY 121 TCGCTTCTGACACACAGGGGGATGATACCATCCGATCCGAGGCTTTCTGGGCAA 180
DB 279 tcgcttcttgacacacaggggagattgatccalcgltacccgtggagcttcttgccaa 338
QY 181 TCTGCAAGCAATTCCTCCAGAGTGATCACTGCACCCACAGATGATTAAGTTGAC 240
DB 339 tctgcaagcaatctccctccaaagtgatcaactctccaccacaagtatataaattgac 398
QY 241 TTAAGTTTCACTGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 300
DB 399 ttaacttcaactcgacttcgcacagactatttctgtaaggtccaggaagcttggtt 458
QY 301 ACCGATTCAGTTAGAACAAAAGCAACACAGAAATCTTGACTACTATATATAA 360
DB 459 agcgattcagttagaaacaaaagcaacaaagaaatcttgacctactatataataa 518
QY 361 GGTTCAGATGTTAATGGAAGTATGGAAGTGAAGCAGAGCTCAAAATCTATGTTAA 420
DB 519 ggttcagatgtttaaagtaactatggaatgcagacagactcaaaactactatgttaa 578
QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTGCTGGCTGGAANTGCTCAGGACC 480
DB 579 agacctcaataattaaagtttaacctagctctgctgctgagatgcttcaagccacc 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGAACGCCCAAGCCGCCAAACTTGCTTT 540
DB 639 aaaccaataatgacccctattcacatcagaacagcccaagccgcgaacttgctt 698
QY 541 ATCTGAATGAAAAATCAGAGCTACATCTCTGTAACAGTATGAAAGCAGTCAATAC 600
DB 699 atctgaaatgaaaaatcaaggttacaatctctctgaacagtattgaaaaagcgtcaatc 758
QY 601 ACCAATTAATGATGACTAACAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGA 660
DB 759 accaatctatgtgactacaagaatctcaatcagcaagttaattaccctgcttaac 818
QY 661 TAATTAACCTCAAGAGATCATCAATCAAGTGAAGAAGAAAGGCTATTAACCTACTCAC 720
DB 819 taattacctcaaggaagtcatcaatcaaggttgaagaagaactggtctataaccttcaac 878

QY 721 AACTGGAGATGATGTTCTACAAATAGTAGACCAAGAGCTCAAAACATCTGTGGATAT 780
DB 879 tactggagatgattgttaacaacaattgtagccaagaagctcaaaaactctgtggatat 938
QY 781 TTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATTTGCAAGTGTCTTACAT 840
DB 939 ctacaactcagatcaatacagctcttaccctcagatgtattgcaagtcgactcagct 998
QY 841 TGTGTATGTTTCAACGTAAGTCAATGTCACGTAAGAGACGCGATCAGTCAAGTAA 900
DB 999 cgtatgttcttaacaatgtaaaagtaacgcacaacttggctgcgtacaaagcaagtaa 1058
QY 901 TGTTCCTTCGAAATTAACGTAAGTCAAGACGATGAACCAACCCGATGGGATCAACTATGA 960
DB 1059 tgttctatcgttaccacacacagcgctgaagaacaaatcgtagctggggtatcaatgaa 1118
QY 961 ACCGATCAGACATATGCTCTGCTGCTGAGATGAGGTGTCTAGATTTCAACTGCTACTAT 1020
DB 1119 accaatcactgactatgctcccgctttagaataatgagcttactactactctat 1178
QY 1021 CGTTACGATGAGCCCTATTAACCTCGGACCAATTAACCTGTTATTAACGTTGATG 1080
DB 1179 tgltaacttgtaaacataatacaatccagatgctcttcaacatacagaaatgtaacagc 1238
QY 1081 GGGCTACTTTGGCAACATCACCCTTGCAATACGCCCTGCAACATCGGAAAGCTCCACAC 1140
DB 1239 tgtctacttggtaaaataatacaatccagatgctcttcaacatacagaaatgtaacagc 1298
QY 1141 CGTGAACCTTAACACAGGTGAGACTACACCGCCCAAGACTTTCTTAATGTCTAGC 1200
DB 1299 cgtgagacttgaataagtcgctctagatagagctaaacctcttcaatgcttgg 1358
QY 1201 AATGCACTACCAAGATTAATCACTACCAATGCAATTAACCAATCAACCAAGATTCACA 1260
DB 1359 talcgacttccaaagatgacatgacaaagcaatccaaatlaaaoaacgaaaccaa 1418
QY 1261 CAAAAAATATGAGACCAAGTATGTAAGAGTGGCTGCTGCTTACGCTTTCGAAATG 1320
DB 1419 caaaaaatctgtgcaagtagtaagtaaaatgctgctgctcagctcttgcctaaatg 1478
QY 1321 TGAACCTTAATTAACCAATGATATTCATTAAGTGTCTTATGATGGAAGTGA 1380
DB 1479 tggatttctcaacaacaaatglatatacaataaaatcgcttcttagtgaatgtagcaaaa 1538
QY 1381 AGAGTTCTTATGTCGGAATCTCGGCCATGGAAGAAAGACAGCTATATGATACCGA 1440
DB 1539 agaatttctgactgtgtaacagactatgaagaagactactgctcatatgactga 1598
QY 1441 CATGATGAACAGTCTTGACTTATGGAACCTGAGCAAAATGCTTATGCTTGGCTCC 1500
DB 1599 aatgataaactgttcttaacttaacagacagagctgtagtgccttaactcaatgcttcc 1658
QY 1501 TCAAGCTGTTAAACAGACGACTCTTAATATACAGACGAAATTTGAACACACATCA 1560
DB 1659 acaagcaggtlaagacaggtactcttaactatactcagaaagaatgaaatgatacaaa 1718
QY 1561 GACCTGCAATTTGTAAGCCTGATGAATATTTGCTGCTTATGCGCTTATGCGCTTAA 1620
DB 1719 gaacactgttactgtagccacagatgaatgcttgaaggtatcacacaaatgacaat 1778
QY 1621 GCGTATGATGACAGGCTATTTAACCCTGTCACACACTTGTAGCAATGAGCTTACGGT 1680
DB 1779 ggcgttggtaagaagataactcaaatcgttcaactcaacatcagagatggttctcgt 1838
QY 1681 CGCTGCAAAAGTTTACCGCTGTATGATGACCTACTGTTGTAAGAGCAATCCAGAGA 1740
DB 1839 tgcgtgtaagctatcgttcaactgaataactcttctgtaagttgacccaacgtaga 1898
QY 1741 TTGGAATATACAGAGGCGCTCTACAGAAATGAGCAATTCGATTTAAATAATGTCGTC 1800
DB 1899 tlgacaatctgcagatggtctgtatagaaaatgagaatctgatttcaaaaatcgtgtcgt 1958

QY 1801 TTCTACGTGGAACTGACCTGCTCCACAAACACCCCATCATGTAAGTTCAAGCTCATC 1860
 DB 1959 ttctacgtggagctcaccgcgtcccaacaacccccatcaactgaagttaagctcacc 2018
 QY 1861 ATCAATTAAGTCTCACTTCACTGCTTCACTCACTCACTCACTCACTCACTCACTCACT 1920
 DB 2019 atcagatagttcaactcaactcaactcaactcaactcaactcaactcaactcaactcaact 2078
 QY 1921 TACCAATCTTACCAATTAAGCAACAAATCAATCAATCAATCAATCAATCAATCAATCA 1980
 DB 2079 taaccatcccaataaataacgaacaatacaatacaactccctgatacaacaatacga 2138
 QY 1981 TCCTCACCAGCACAACCA 1999
 DB 2139 tctcaccagcacaacca 2157

RESULT 9
 AAH02147
 ID AAH02147 standard; DNA; 2157 BP.
 AC AAH02147;
 DT 24-JUL-2001 (first entry)
 XX
 DE Streptococcus pneumoniae nucleotide sequence SRO ID NO:2140.
 XX
 KW Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitical;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200123604-A2.
 XX
 PD 05-APR-2001.
 XX
 PE 28-SEP-2000; 2000WO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 XX
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI: 2001-245006/25.
 XX
 XX Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitical species in a test sample -
 XX
 PS Disclosure: Page 1508-1509; 1580pp; English.
 XX
 CC The present invention describes a method for generating a repository of
 CC nucleic acids of tuf, fns, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (1)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.

CC (1) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SO Sequence 2157 BP; 696 A; 494 C; 405 G; 562 T; 0 other;

Query Match 80.4%; Score 1607; DB 22; Length 2157;
 Best Local Similarity 87.7%; Pred. No. 0;
 Matches 1754; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATTAATCAACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 DB 159 taaatctacgacaataaatacaactcatctgctgctgctgctgctgctgctgctgctgctgct 218
 QY 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGGCAATCGTTTATCGAAGACCA 120
 DB 219 tgcccaagctaatgatattccacagatttggttaaggaatcgtttcatcgaagacca 278
 QY 121 TGGCTTCTTCGACACACAGGGGATTTGATTCGATTCGCTGCGGACCTTCTTGCGCAA 180
 DB 279 tggcttcttcgacacacaggggattgattccatccgtatccgggagcttcttcgcaaa 338
 QY 181 TCGCAAGCAATTCCTCCCAAGGTGATGATCACTGCACTCCCAAGGTTATTAAGTTGAC 240
 DB 339 tctgcaagtaattccctcccaaggtggaatcaactccacaagtgattaaagtgcac 398
 QY 241 TTACTTTTCACTGCACTTCGACAGACTATTTCTCTGAAGGCTCAGAAAGCTTGTT 300
 DB 399 ttactttcaactcgacttcgacagactatctctgtaaggtcaggaagcttggtt 458
 QY 301 AGCGATTGCTTGAAGCAAAAGCAACGCAAGCAAAATCTTGACTACTATATTAATA 360
 DB 459 agcgattgcttgaagcaaaagcaacgcaagcaaaatcttgactactatataataa 518
 QY 361 GGGCTACATGCTGAATGGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 DB 519 ggcctacatgctgaatggaaatgatatgatatgatatgatatgatatgatatgatatgatat 578
 QY 421 AGACCTCAATTAATTAAGTTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 579 agacctcaatatttaagtttaactctgctgctgctgctgctgctgctgctgctgctgctgct 638
 QY 481 AAACCAATATGACCCCTATTCACATCCAGACGAGCCCAAGACCGCGAACTTGGCTTT 540
 DB 639 aaaccaatatgacccctatttcacatccagaagcagccagcagcagcagcagcagcagcagc 698
 QY 541 ATCTGAATGAAGAAATCAAGGCTACATCTCTGCTGAACGATGANAAGCAAGCTCAATAC 600
 DB 699 atctgaaatgaagaaatcaaggctacatctctgctgaacgattganaagcagctcaatc 758
 QY 601 ACCAATTAAGTATGATGCTCAAAAGTCTCAAAAGTCTCAAAAGTCTCAAAAGTCTCAAA 660
 DB 759 accaatattagttgactcaaaagctcaaaagctcaaaagctcaaaagctcaaaagctcaaaag 818
 QY 661 TAATTAAGTCAAGAGATGATCAATCAAGTGAAGAAAGCAAGCTTATACCTACTAC 720
 DB 819 taattaaagtcgaagatgataatcaatgaagaaagcaagcttattaccctactaac 878
 QY 721 AACTGGATGATGATGCTTACCAAAAGTGTAGCAAAAGTGTAGCAAAAGTGTAGCAAA 780
 DB 879 taactggatgattgatttacaacaagtgtagcaaaagtgtagcaaaagtgtagcaaaagtgtag 938
 QY 781 TTACATATACAGAGATATGATGCTTATCCAGACGATGATGATGATGATGATGATGATGAT 840
 DB 939 ctacaactcgttcaataagcttcttacccttgagatgatttgcaagtcgactcaaggt 998

QY	841	TCGTTGATGTTTCTPAACGGCTAAACACTCATTTGCCAGCTAGGAGGACGCCATCATGCTCAAGTAA	900
Db	999	cgtagagcttccaatggttaaaagctatcgccaaactcttgagctctgaccagaagctaa	1058
QY	901	TGTTTTCCTGGGAATTAACCAAGCAGTATAGAAACCAACCCGACTGGGATCAACTATGAA	960
Db	1059	cgcttcatttggtaaccacaagctgtagaaccatctgtagcgggttcgtatgaa	1118
QY	961	ACCGATCACAGACTATGCTCTCTGCTTGGAGTACGGTGTCTACGATTCACCTGCTACTAT	1020
Db	1119	accaataccgatatgacacctgcataagatacggcttgcttatctccactgcgaactat	1178
QY	1021	CGTTACAGATGAGCCCTTAATCTACCTGGGAGAAATATCTCTGTTTAACTGGGATAG	1080
Db	1179	ggttaatgatatctctataactatccgcggaataagcaaacctgtctcaactcgtgtag	1238
QY	1081	GGGCTACTTTGGCAATCATCTTGCATATACGCCCTGCACAAATCGAAGAACGTCCAGC	1140
Db	1239	agcaatctcgttatatctactcttcgaataatgctcttcacatacgaatgctacagc	1298
QY	1141	CGTGGAAACTCTAAACAAGGTGGGACTCCACCGCGCAAGACTTTTCTCAATAGGCTTAG	1200
Db	1299	cgltgagacttgtaataagctcggtccagataagataaaccttcctaaagctcgtg	1358
QY	1201	AATCGACTACCCAAATTTACTACTCAATATGCGATTTTGAATTAACACACGATTCAGA	1260
Db	1359	tatcgactatccaaagcatgcatattgcaaacgcattccaagtaatacaacagatctaa	1418
QY	1261	CAAAAAATATGGAGACAGTATGTAAGAAAGATGCGTCTGCTTACGCTGGCTTTGGCAATGG	1320
Db	1419	taaaacaatacggagcaagtagtgaaaaaaltgctgtgcttatgcttgctcttgcacaag	1478
QY	1321	TGGAACTACTATTAACCAATGTATATCCATTAAGTCGCTTATAGTATGGAGTGAATA	1380
Db	1479	tggaattaccacaacaccatctgatatcaataaagtcgcttcagtgagcgttgatgaaaaa	1538
QY	1381	AGAGTTCTCTAATGTGGAACCTGTGCCATGAAGAAGACAGACGCTTATATGATGACCGA	1440
Db	1539	agaattctcagatgtagtgcacgagctatgaaagaaacaaactgcttacaatgatagcga	1598
QY	1441	CATCATGAAAAAGCTGTTGACTTATGGAATCGAGCAAAATGGCTTATGGCTGGCTCC	1500
Db	1599	aatgataaaaaactcgtctctgcatagacggaactcgtgtgtagcctatcccatggttgc	1658
QY	1501	TCAGGCTGTGTAACAAGACACTCTTAACATACAGACGAGAAATTGAAAAACCATCAA	1560
Db	1659	gcaagctcgttaagacaagctactcttcaactacaagtagtgaaattgtaaaaaacaacaa	1718
QY	1561	GACCTCCAAATTTGTATACCACTGATTAATTTGCTGTGGCTATACGGTAAATATTCAAT	1620
Db	1719	gaacaactggtatgtagctccagaatgaaltgttctgttatactcgttaagatctcat	1778
QY	1621	GGCTGTATGGACAGGCAATCTPAACCGTGTGACACACTTGTAGGCAATGGCCTTACGCT	1680
Db	1779	ggtcgtatgtaacaggttatctcgatatcgtttaactcctacgtcttgagtggtttccctagt	1838
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGTATGACTTAACCTGTGTGAGGAGACCAATCCAGAAGA	1740
Db	1839	tgcagctaaagtttatcgtccataatgataacgtatctatcaagaagaatactcatccagaaga	1898
QY	1741	TTTGAATATATCCAGAGGGGCTCTACAGAAATGGAGAAATTCGTAATTAATAATGGGCTCG	1800
Db	1899	ctggagcgaatgcacgaaggaactttccagaaacggggaatttgcattccaataaayagagctcg	1958
QY	1801	TTTACCTGGAACTCACTGCTTCACACAAACCCCATCACTGAAGTTCAAGTCTATC	1860
Db	1959	tctcactggtgagctcaccctgcctccacaacacccccaactcaatgaagtccaagctatc	2018
QY	1861	ATCAGATATGTTCAACTTACAGCTTATAGCTCAACACACTCAACACCAATATATAGTACAC	1920
Db	2019	atccagatagttcaacttcaacagctctagcttcaacacactccaagcaacaaataatagtaacac	2078

[illegible]

RESULT 10

ID AAZ35952 standard; DNA; 1260 BP.

AC AAZ35952;

DT 07-FEB-2000 (first entry)

DE Streptococcus pneumoniae pbp1A T_{ER} isolate n) nucleotide sequence.

KW Streptococcus pneumoniae; penicillin binding protein; pbb2B; pbb1A;

KW detection; identification; pneumococcal meningitis; ss.

Streptococcus pneumoniae.

PN ZA9807024-A.

PD 28-APR-1999.

PF 05-AUG-1998; 98ZA-0007024.

PR 01-AUG-1997; 97ZA-0006886.

PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.

PA (MEDI-) MEDICAL RES COUNCIL.

PI Klugman KP, Smith AM, Du Plessis M;

DR WPI; 1999-601770/51

PT polymerase chain reaction assays for detecting *Streptococcus pneumoniae*

XX
XX
2
.
4
1
.
.
2
2
1
1
1
1
1

XX
XX
XX

CC detecting an antibiotic resistant strain of *Streptococcus pneumoniae*

CC and the pbp1A gene. The products and methods can be used for detecting

used for simultaneously diagnosing pneumococcal meningitis and

CC The methods can be used for detecting *S. pneumoniae* strains resistant

CC The assays can be adapted to detect other pathogens causing meningitis

S. pneumoniae with a minimum inhibitory concentration (MIC) of

CC a 224 bp product. The present sequence represents a *Streptococcus*

CC sequence from the present invention.

Sequence 1260 BP; 407 A; 292 C; 259 G; 302 T; 0 other;

CONFIDENTIAL

Best Local Similarity	99.68;	Pred. No. 0;
Watchdog 1055	0;	Watchdog 5
Concomitance	0;	Indolence
Concomitance	0;	Concomitance

EZE 631

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

```
OY 635 GCAAGTAATTAACCCGCTTACATGATTAATTACCTCAGAAGATCATCATCAAGTTGAA 694
    |||||||
Db 61 gcaagtaattacccctgcttacaatgataactcaagaagatcaatcaactgagaa 120

OY 695 GAAGAAACAGGCTTAATACCTACTCACAACGTGGATGATGCTCTACACAATATGAGACCAA 754
    |||||||
Db 121 gaagaaacaggtctataactactcaactcagtgatgagtgctctacacaaagttagaccaa 180

OY 755 GAAGCTCAAAAACATCTGTGGGATATTTACAATACAGCAATACGTTGCTATTCAGAC 814
    |||||||
Db 181 gaagctcaaaaacatctgtggatatactacaatacagaacatagcttgcctatccagac 240

OY 815 GATGATATGCAAGTGGCTTACATGTTGATGTTTCTAACGGGAAAGTCAATGGCCAG 874
    |||||||
Db 241 gatgatataagtcagctcgtcttacaatctgtatgtcttcaacggtlaaagltcatctgcag 300

OY 875 CTAGAGACGCCATCATCAAGTAATGTTTCTTCGGAATTAACCAAGAGTAGAACA 934
    |||||||
Db 301 cttagagacgcgcatacgtcaagtaatgttccctcggaattaaaccaagcagtgaacaa 360

OY 935 AACCGGCACTGGGATCACTATGAAACCGATCACAGACTATGCTCTGCTTGGAGTAG 994
    |||||||
Db 361 aaccggcaactgggatacaactatgaaccga tcaagactatgctcctccttagagtagc 420

OY 995 GGTTCTACGATTCACGCTACTATGCTTCACGATGAGCCCTTAATACCTACCGGGACA 1054
    |||||||
Db 421 ggttctacgattccaactgctcaactatcgttcaacgtgagcccttaactaccctgggagca 480

OY 1055 AATACTCCTGTTTAACTAGGAGATAGGGGCTACTTGGCAACATCACCTTCAATACGCC 1114
    |||||||
Db 481 aatactcctgttataactcgtgataaggggtactcttgccaacatcaccttgcaataagcc 540

OY 1115 CTGCAACAAATGCGCAACAGTCCGACCGTGGAACCTTAACAAGGTGGACTCAACCGC 1174
    |||||||
Db 541 ctgcaacaatcgcgaaacgtccacgcgtggaactcaaacgaagtcgacctaaccgc 600

OY 1175 GCCAAGACTTTCCTAAATGGCTGTGAATGCATACCACTATTCCTACTCTCAATGCC 1234
    |||||||
Db 601 gccaaagactctccaaataggctaggaatcgatacccaaglatctactactcaaatgccc 660

OY 1235 ATTTCAGTAAACACACAGCAATCAGACAAAATAATGAGCAAGTAGAGAAAGATGGCT 1294
    |||||||
Db 661 attcaagtaaacacacacacgaatcagacaaaataatgagcaagtaggaagaatgagct 720

OY 1295 GCTGCTTAACGCTGCCCTTGCATAATGTGGAACCTTACTATAAACCAATGTATATCAATAA 1354
    |||||||
Db 721 gctgcttaacgctgcctctgcaaatggtggaacttactataaacaatgttatatacataaa 780

OY 1355 GTGCTCTTAACTGATGGAGTGAAGAAAGAGTTCTCTAATGTCGAACTCGTCCCATGAG 1414
    |||||||
Db 781 gtgctctttagtgaaggagaaagagttccttaatgctggaactcgtgcatatgag 840

OY 1415 GAAACGACAGCCTTATGATGACGACATGATGAAACAGTCTGACTTATGGAAGTGA 1474
    |||||||
Db 841 gaaacgacagcctataatgataigaccgacatgataaacaagcttgagttatggaactgga 900

OY 1475 CGAATGCTATCTTGTGCTGCCCTCAGGCTGTTAAAGAGAGAACCTTACTATTAACA 1534
    |||||||
Db 901 cgaatgctatcttctgtgctccctcgaagcgtgtaaaaaagaaacctcaactataca 960

OY 1535 GAGCAGGAAATTTGAAGAACCAATCAAGACCTCAATTTTGAAGACCTGATGAGACTATTT 1594
    |||||||
Db 961 gagcaggaataattgaagaacacacatacagccctcaatttgaagaccctggaagacattt 1020

OY 1595 GCTGCTATACGCGTAAATATTAATGCTGATGAGACGCTATTCTAACGCTGTGACA 1654
    |||||||
Db 1021 gctgctatacagcgtaataatataatgctgtatgagacagctatcttaaccgcttgaca 1080

OY 1655 CCACTTGTAGGCAATGGCCTTAGGGTGCCTGCCAAAGTTTAACCGCTCATGATGACCTAC 1714
    |||||||
Db 1081 ccaacttgtagcaaatgagcttaacgctcgtcgcaaaagttacgcgtcatatgatgacctac 1140
```

```
OY 1715 CTGCTGAGAGGACAGCAATCCAGAAGATTGAAATATACAGAGGGGCTCTACAGAAATGGA 1774
    |||||||
Db 1141 ctgctgtagaagaaagcaatcccgagagattggaatatataccagagggctctacaagaatgga 1200

OY 1775 GAATTCGATATTTAAATAATGTCGTTCTCTACGTGGAACCTGACCTGCACAAACACC 1834
    |||||||
Db 1201 gaattcgatttaaaaatagtgctgctgtctacgtggaactcacctgcctccacaacaccc 1260

RESULT 11
AAZ35950
ID AAZ35950 standard; DNA; 1260 BP.
XX
AC AAZ35950;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1a T8R isolate 1) nucleotide sequence.
XX
KW Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1a;
KW transpeptidase encoding region; T8R; antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 98ZA-0007024.
XX
PR 01-AUG-1997; 97ZA-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UYMI-) UNIV WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
DR WPI: 1999-601770/51.
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
XX
PS Claim 11: Fig 4; 63pp; English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1a transpeptidase encoding region (T8R) isolate nucleotide
CC sequence from the present invention.
XX
S0 Sequence 1260 BP; 407 A; 293 C; 258 G; 302 T; 0 other;

Query Match 62.5%; Score 1248.8; DB 20; Length 1260;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 575 GAACAGTATGAGAAAGCAGTCAATACCAATTAAGTGTGAGTACAAAGTCTCAAAATCA 634
    |||||||
Db 1 gaacagtatgagaaagcagtcataatacaccattactatgtagactacaagaatcctaataca 60
```

OY	635	GCAGTATTACCGCTTGCATGAGTAATTAATTAAGGAGTCAATCAATCAAGTAA	694
Db	61	gcaagtaatactccctgcgtctaaagtgataaattlaccctaagaagtaactaaatgaatlgaa	120
OY	695	GAAGAAACAGGCTTTAACTACTACAACTGGGATGGATGCTACACAATTTAAACCAA	754
Db	121	aaggaacaaggctataaaccctgcgtctacaactcggatgtaactgtctacacaaatgtagaccaa	180
OY	755	GAAGCTCAAAAACATCTGTGGGATATTTACAATACAGAGGAATACGTTGCCATATCCAGAC	814
OY	815	GATGAATTCGAAGTCGGCTTTACCATTTGTGATGTTTCTAACAGGTAAATCATTCGCCAG	874
Db	241	gatgaattgcgaagtcgcgtctctacatctggtatgattcttcaagtaagaatctatgcccag	300
OY	875	CTAGAGCAGCCCTCATCGATCAATATGTTTCTTGGGAATTAACCAAGCACTAGACAACA	934
Db	301	ctaggaagacagccctacagtcgaatgaatgatttccctcgtgaattaaaccaagcagtaagaaca	360
OY	935	AACCGGCACTGGGGATACACATGAAACCGATACAGACTATGTCCTCCCTGGAGTAC	994
Db	361	aaccggcactcggggatcacactaagaaacgataccagactatgcttcgccttggatgac	420
OY	995	GGTGTCTACGATTCAGCTGCTACTATGTTTCAACAGTAGAGCCCTATTAACATCCCTGGAGCA	1054
Db	421	ggtgtctcagatctcaactgctactatcgtctcagatgtagccctataactacactccttggaga	480
OY	1055	AATACTCCTGTTTAACTGGGATAGGGGCTACTTTGGCAACATCACTTGCATATAGCC	1114
Db	481	aatctcctglttataactcgtggaataggggtactcttggacaatacaactcttgaataagcc	540
OY	1115	CTGCAACAAATGGCGAAACGTCGCCAGCCGTGGAACCTCTAAACAGGTGCGATCAACGC	1174
Db	541	ctgcaacaatcgtcgaaacgctccagccgcgtggaaccttaacaagtgctggaactcaacgc	600
OY	1175	GCCAAAGACTTTCTTAAATGGCTCTTAGGAATGCACTACCCAAATTTACTACTCAATAGCC	1234
Db	601	gccaaagactttccttaaaatgctcgtgaatcgactacccaagatatacctaactcaaaatgcc	660
OY	1235	ATTTCAGTAAACACAAACCGAATCAGACAAAAAATTTGAGCAAGTAGTAAAGATGGCT	1294
Db	661	atttcaagtaacaacaaccggaatcagacaaaataatgaggaacagtcagtgaaagaatgctc	720
OY	1295	GCTGCTTACGCTGCCCTTGGCAAAATGGTGAACCTTACTATTAACCAATGATATCCATAA	1354
Db	721	gctgcttacgctgccttggcaaatgtygnaacttactataaacaatgataatccataaa	780
OY	1355	GTCGCTTTAGTGAATGGAGCTGGAAGAAAGTCTCTAATGTGGAACTGTGCCATAGAG	1414
Db	781	gtcgcttctaaatgtagtgagtgagtaaaaagatctctcctaaatgctggaactcgtgtccatgaag	840
OY	1415	GAACAGCAGGCTTTATGATGAGACCGCATGATGAAAAACAGTCTTGACTTATGGAAGTGA	1474
Db	841	gaacaacagcctataatgatgacgcagacatgatagaanaacagctctgagtatgtaagacttga	900
OY	1475	CGAAATGCTATCTTGGCTTGCTCCTCAGGCTGGTAAACAGAGCACTTAAGTATATCA	1534
Db	901	cgaatgtcctatctctgcttgcgtccctcagcgtctgttaaaacaggaaccttcaactataca	960
OY	1535	GACGAGGAATTTGAAAACCAATCAAGACCTTCATTTGTGACACGTGATGAATCAATATT	1594
Db	961	gacgaggaatatgtaaaacacacatacgaagcccttccaattgtatgacacgtgaacgaactatct	1020
OY	1595	GCTGGCTATACGCGTTAAATTTCAATGGCTGTATGTGACAGGCTATTTCAACGCTTGACA	1654
Db	1021	gctgctatacgcgttaaaataatcaatagtcgtatgtgacagcgtctattcttaacgcgtctgaca	1080
OY	1655	CCAATTGTAGGCAATGGCCTTAAGGTCGCTGCCAAAGTTTACCCTCTATATGATACCTAC	1714
Db	1081	ccaatctgtagcaatctgcttcaagtcgtcgtcgcacaaagtttaccgcctctatgatgtgacctac	1140

Query	Subject	Score	Length	Start	End	Ident	Match	Mismatch	Indel	Gap
1715	CNGTCGAGGAAGCAATCCAGATTGGATATATCCAGAGGGGCTCTACAGAAATGGA	1248.8	1260	1	1260	100.0	0	0	0	0
1141	ctgtctgaagaagcaatccagagatctggaatataccagagggctctacagaaatgga	1200	1260	1	1260	100.0	0	0	0	0
1775	CAATTTCGATTTAAATAATGCTGCTGTTACGTGCAATCTACCTCTCCACAAACCC	1834	1260	1	1260	100.0	0	0	0	0
1201	gaattcgtattttaaaatggtgctgtcttaagtgaactcaactgcttcacacaacc	1260	1260	1	1260	100.0	0	0	0	0
AAZ35951	standard; DNA; 1260 BP.									
AAZ35951	AAZ35951									
07-FEB-2000	(first entry)									
Streptococcus pneumoniae pbp1A TER isolate m)	nucleotide sequence.									
Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A; transpeptidase encoding region; TER; antibiotic resistance; diagnosis; detection; identification; pneumococcal meningitis; ss.										
Streptococcus pneumoniae.										
ZA9607024-A.										
28-APR-1999.										
05-AUG-1998;	982A-0007024.									
01-AUG-1997;	972A-0006886.									
(SAME-) SOUTH AFRICAN INST MEDICAL RES. (UWI-) UNIV WITWATERSRAND. (MEDI-) MEDICAL RES COUNCIL.										
Klugman KP, Smith AM, Du Plessis M; WPI; 1999-601770/51.										
Polymerase chain reaction assays for detecting Streptococcus pneumoniae useful for the diagnosis of pneumococcal meningitis										
Claim 11; Fig 4; 63pp; English.										
A polymerase chain reaction (PCR) assays have been developed for detecting an antibiotic resistant strain of Streptococcus pneumoniae using primers based on the penicillin binding protein 2B (pdp2B) gene and the pbp1A gene. The products and methods can be used for detecting S. pneumoniae, particularly antibiotic-resistant strains. They can be used for simultaneously diagnosing pneumococcal meningitis and identifying any antibiotic-resistant S. pneumoniae strains in a sample. The methods can be used for detecting S. pneumoniae strains resistant to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin. The assays can be adapted to detect other pathogens causing meningitis. S. pneumoniae can be used to detect an antibiotic resistant strain of S. pneumoniae with a minimum inhibitory concentration (MIC) of 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and a 224 bp product. The present sequence represents a Streptococcus pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide sequence from the present invention.										
Sequence 1260 BP; 407 A; 293 C; 258 G; 302 T; 0 other.										
Query Match	62.5%	Score 1248.8	DB 20	Length 1260						
Best Local Similarity	99.4%	Pred. No. 0								
Matches 1253	Conservative 0	Mismatches 7	Indels 0	Gaps 0						
575	GACACGATAGAAAGCAGCTCATATACCAATTAATGATGAGCTACAAAGTCTCAATCA	634	1260	1	1260	100.0	0	0	0	0
1	gaacagatgaaagacgctcaatcaccaatattactgattgagctacaagtctcaatca	60	1260	1	1260	100.0	0	0	0	0

QY 635 GCAAGTAATTACCCCTGCTTACATGATGAATTAATTAACCTCAAGAACTCATCAATGAATTGAA 694
|||||
Db 61 gcaagtaattaccctgcttaccatgataatctaccctaaggaagtcataccaatgaa 120
QY 695 GAAGAAACAGGCTATTAACCTACTACAACTGGATGATGTTCTACACAATAATGAGACAA 754
|||||
Db 121 gaagaaacaggctataaccctaccacaactcggatgagatgctctaccacaatgttagacca 180
QY 755 GAAGCTCAAAAACATCTGTGGATATTTACATACAGACAAATACGTTGGCTATCCAGAC 814
|||||
Db 181 gaagctcaaaaaacatctgaggatatttaacaacagaaatagcttgcctaccagac 240
QY 815 GATGAATTCGAAGTCGCTTCTACCAATGTTGATGTTCTTAACGGTAAAGTCATTTGCCAG 874
|||||
Db 241 gatgaatcgaaagtcgcttaccacatgtgatgtcttcaagcgtaaagcatcgccag 300
QY 875 CTAGAGACAGCCATCAGTCAAGTATGTTTCTTGGGAATTAACCAAGCATAGAAACA 934
|||||
Db 301 ctagagacagccatcagtaagtaattctcttcggaatlaacaaagcagtagaaca 360
QY 935 AACCGGACTGGGATCACTATGAAACGATCAGACATATGCTGCTGCTGGAGTAC 994
|||||
Db 361 aacccgactggggatcaactatgaaacccgatacagactatgcttcgcttggagtac 420
QY 995 GGTGTCTACGATTCGAAGTCTACTATGTTTACAGATGAGCCCTATTAACCTGGGACA 1054
|||||
Db 421 ggtgtctacgatctcaactcgtactatcgttcaagatgagccctatacccttggaca 480
QY 1055 AATACCTCTGTTTATTAACAGGGATAGGGGCTACTTTGGCAACTCATCTGGCAATACGCC 1114
|||||
Db 481 aataccctgttataacacggatagggatctacttggacaacatacccttgcatacgc 540
QY 1115 CTGCAACAATCGGAAAGCTCCAGCGGTGAAACTCTAAACAGGTGGAGCTCAACCGC 1174
|||||
Db 541 ctgcaacaatcggaaacgtccacgcgtggaaaccttaacaagctcggactcaaccgc 600
QY 1175 GCCAAGACTTTCTTAATGTTCTAGGAATCGACCTACCAAGTATTCATCTCAAAATGCC 1234
|||||
Db 601 gccaaagacttctctaattgctcgaatcgaactaccacaagtaattcaactcaaatgccc 660
QY 1235 AATTCAGTAACACACACCAATCAGACCAAAAATATGAGCAAGTATGTAAGAAAGATGCGT 1294
|||||
Db 661 atttcaagtaacacaacacgaatcgaacaaaataatgagcaagtagtgaagaagcgtc 720
QY 1295 GCTGCTTACGCTGCTTCTCAAAATGTTGAGACTTACTATAAACCAATGATATCCATAA 1354
|||||
Db 721 gctgcttaacgctgcttgcataatggttgaaacttaactataaccatgatatccataaa 780
QY 1355 GTGCTCTTATGATGAGGAGTGAAGAAAGATCTCTTAATGTGGAACCTGTCGCATGAAG 1414
|||||
Db 781 gtctctttagtgatggagtgaaaagagctctctaactgctgaactcgtgcataagaa 840
QY 1415 GAAAGACAGCCTATATGATGACCCAGATGATGAAGAAAGCTCTTGACTTATGGAAGTGA 1474
|||||
Db 841 gaaagacacgacctataatgtagccacatgataaagagctctgtagtataatggaactgga 900
QY 1475 CAAAATGCTATCTTGTGCTGGCTCCCTGAGGTGGTAAACAGGAACCTCTAATCTTACA 1534
|||||
Db 901 cgaatgctactcctgcttgcctccctcagcgtgtaaaaacaggaacctctaactataca 960
QY 1535 GACGAGGAATTAAGAAACACCATCAAGACCTCTCAATTTGTAGCACCTGATGAATATTT 1594
|||||
Db 961 gacgaggaatgtgaaaacacacatcaagacctctcaattgttagcactcgtatgaactatt 1020
QY 1595 GCTGGCTATACCGCTAATATTTCAATGCTGTATGACAGGCTATTTTACCGCTGACA 1654
|||||
Db 1021 gctggctatacgcgtataatcattcaatgctgtatgacagagctatcttaaccgctcgaca 1080
QY 1655 CCACCTGTATGGAATGAGCGCTTACGGGCTGCGCAAAAGTTTACCGCTCTGATGAGCTTAC 1714
|||||
Db 1081 ccactctgtagcaatcgcccttaacggtcgcgcgcaaaagtttaccgctctatgtgacctac 1140

QY 1715 CTGCTGAAAGAACCAATCCAGAGATTGGAATATACAGAGGGGCTCTACAGAAATGCA 1774
|||||
Db 1141 ctgctgaaaggaagcaatccacagagattggaatataccagaaggcgtctacagaatgga 1200
QY 1775 GAATTTGCTATTTAAATGTTGCTGTTCTACTGGAACCTGACTGCTCCACAAACACCC 1834
|||||
Db 1201 gaattcgtattataaatggtcgtcttactcgttgagagctacacctcctcacaacaacc 1260

RESULT 14
AAH02067
ID AAH02067 standard; DNA; 1249 BP.
XX
AC AAH02067;
XX
DT 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2060.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW Identification; algal; archaeal; bacterial; fungal; parasitical;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200123604-A2.
XX
PD 05-APR-2001.
XX
PE 28-SEP-2000; 2000WO-CA01150.
XX
PF 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
DR WPI; 2001-245006/25.
XX
PT Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitcal species in a test sample -
XX
PS Claim 27; Page 1473; 1580pp; English.

The present invention describes a method for generating a repertory of nucleic acids of tuf, fus, atpD and/or recA genes from which probes and/or primers are derived. The method comprises amplifying the nucleic acids of determined algal, archaeal, bacterial, fungal and parasitcal species with a combination of defined primer pairs. The method can be used for producing probes and/or primers for detecting one or more related microorganisms e.g. algae, archaea, bacteria, fungi and parasites, for universal detection and for specific and ubiquitous detection and identification of an algal, archaeal, bacterial, fungal and parasitcal species, genus, family and group. A nucleic acid (1) obtained using the method of the invention can be used for the universal detection of any bacterium, fungus or parasite in a sample and for the detection of at least one antimicrobial agent resistance gene or at least one toxin gene. hexa nucleic acids are used for the specific and ubiquitous detection and for identification of Streptococcus pneumoniae. (1) can be used to design a therapeutic agent which is effective against microorganisms. Microbial species or genus or family or phylum or group which can be detected include Abiotrophia adiacens, Bordetella sp., Corynebacterium sp., Enterobacteriaceae group, Escherichia coli, Mycobacteriaceae family, Pseudomonads group, Streptococcus sp., Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests provides faster results than substrate specificity tests as results can be determined in an hour and improved accuracy is also achieved.

CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 XX Sequence 1249 BP; 402 A; 302 C; 246 G; 299 T; 0 other;

Query Match 62.4%; Score 1247.4; DB 22; Length 1249;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 720 CAACCTGGGATGATGCTACACAAATGTAGACCAAGAGCTCAAAAATCTGTGGAGTA 779
 Db 1 caactggaatgagatgctacacaaatgtagaccagaagctcaaaaatctgtggagata 60
 QY 780 TTACCAATGACAGCGAATGCTTGGCTATCCAGAGATGAATTTGCAAGTGGCTTACCA 839
 Db 61 ttacacaatacagagaaatgcttgcctacacagaaatgtagaagctgcttccaca 120
 QY 840 TTGTGATGTTCTAAGCGTAAAGTCATTGCGCAGCTAGAGACGCCATCAGTCAAGTA 899
 Db 121 ttgtgattgttctaaacgtaaaagtcatgtgcccagctagagacgcatcagtaagta 180
 QY 900 ATGTTTCCCTTGGGATTTAACAAGCAGTAGAACAACCGGAGCTGGGATCACTATGA 959
 Db 181 atgttcccttcggaatlaacaaagcagtagaaacaaacgcgactgggatacactatga 240
 QY 960 AACCGATCAGACTATGCTGCTGCTTGGAGTACGCTGTACGATTTCACTGCTACTA 1019
 Db 241 aaccgatacagactatgcttccctgcttggagtaagtgctacgactaacagctacta 300
 QY 1020 TCGTTCACGATGAGCCCTTAACTACCTCGGACAAATCTCCTGTTTAACTGGGATA 1079
 Db 301 tgcgtcacgatagagcccttaactaacctcgggacaataactcctgtttataactggata 360
 QY 1080 GGGGCTACTTGGCAACATCACCCTTGCATACGCTTGCACAAATCGCAACGCTCCAG 1139
 Db 361 ggggctacttggcaacatcaccttgcatacagccctgcgaacatcggaagctccag 420
 QY 1140 CGGTGAAACTGTAAACAAGGTGCGACTACCGCGCAAGCTTCCCTTAAATGCTCTAG 1199
 Db 421 ccgtggaactctaaacaagtcgagactcaacgcgcaagacttccctaaatgctcag 480
 QY 1200 GAATGACTACCAAGTATTTCACTACTCAATGCCATTTCAAGTAACACACGGAATCAG 1259
 Db 481 gaaatgactacccaagatattcaactactcaaatgcaatcctcaagtaacacacgaatcag 540
 QY 1260 ACAAATAATATGAGCAGTAGTGAAGAGATGCTGCTTACGCTGCTTGAATG 1319
 Db 541 acaaaaataatgagcagtagtgaagagatgctgcttgaagctgcttgccttgcacatg 600
 QY 1320 GTGGAACTTACTATAACCAATGTATATCCATTAAGTCTTGTAGTGGAGTGA 1379
 Db 601 gtggaacttactataaacaatgataatcagcttctttagtggagagtaaa 660
 QY 1380 AAGAGTCTCTAATGTCGGAACCTGCGCATGAAGAGAAAGCAGAGCTTATGATATACCG 1439
 Db 661 aagagcttctaatgctggaactcgtgccatgaagaaacgacagcctataatgataacg 720
 QY 1440 ACATGATGAAACAGTCTTGTACTTATGAACTGGACGAATGCTTATCTTGTGGCTCC 1499
 Db 721 acatgataaagacttctgacttataatgagactggaatgcttcttgccttgcctcc 780
 QY 1500 CTCAGGCTGTAAACAGACCTCTAATATACAGACAGAAATTTGAAACACATCA 1559
 Db 781 ctcagagctgtgtaaaacagacacttaactatacagaagaaatgtaaaacacataa 840
 QY 1560 ACACCTCTCAATTTGTAGACCTGATGACTATTTGCTGGCTATACGCTAAATATTCA 1619
 Db 841 acacctctcaattttagaactgtatgaactatgttgcgtatacagcgtataatcaca 900
 QY 1620 TGGCTGTATGACAGGCTATTTCAACCGTGTACACACCACTTGTAGCAATGGCTTACGG 1679
 Db 901 tggctgtatgacagagctatcttaaccgtctgacacactttaggcaatgagcttaagc 960

QY 1680 TCGCTCCAAAGTTTACCCGCTGATGATGACCTACCTGCTGTGAAGAGCAATCCAGAA 1739
 Db 961 tgcgtccaaagtcttaacgctctatgatacctactgtctgaaagaaatccagaag 1020
 QY 1740 ATTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTAAATAATGCTTC 1799
 Db 1021 attggaatatacagaggggctctacagaatgagaaatctgatttataaatgctcc 1080
 QY 1800 GTTCTACGTGGAATCAGCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGTCAAT 1859
 Db 1081 gtctacgttggagctacactgtctccacaacaccccatcaactgaaatgtaagctcat 1140
 QY 1860 CATCAGATAGTTTCACTTCACTACAGCTAGCTCAACCACTCCAAAGCAAAATATAGTACGA 1919
 Db 1141 catcagatagttcaacttcaacagctcagctcaacccatccaaacacaaataaataagta 1200
 QY 1920 CTACCAATCTTAACAATATACGCAACAAATCAAAATCAACCCCTGATCA 1968
 Db 1201 ctaccatctctaataataatagcgaacaaatcaataacacccctgtatca 1249

RESULT 15
 AA235939
 ID AA235939 standard; DNA; 1260 BP.
 XX
 AC AA235939;
 XX
 DT 07-FEB-2000 (first entry)
 XX
 DE Streptococcus pneumoniae pbp1A TER isolate a) nucleotide sequence.
 XX
 KW Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
 KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
 KW detection; identification; pneumococcal meningitis; ss.
 XX
 OS Streptococcus pneumoniae.
 XX
 FH Location/Qualifiers
 FT 1..1260
 FT CDS
 FT /*tag= a
 FT /*product= "pbp1A TER"
 FT /*note= "no stop codon given"
 XX
 PN ZA9807024-A.
 XX
 PD 28-APR-1999.
 XX
 PF 05-AUG-1998; 982A-0007024.
 XX
 PR 01-AUG-1997; 972A-0006886.
 XX
 PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
 PA (UYWI-) UNIV WITWATERSRAND.
 PA (MED-) MEDICAL RES COUNCIL.
 XX
 PI Klugman KP, Smith AM, Du Plessis M;
 XX
 XX WPI: 1999-601770/51.
 DR P-PSDB; AAY56106.
 DR
 XX
 PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
 PT useful for the diagnosis of pneumococcal meningitis
 XX
 PS Claim 11: Fig 4: 63pp: English.
 XX
 CC A polymerase chain reaction (PCR) assays have been developed for
 CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
 CC using primers based on the penicillin binding protein 2B (pbp2B) gene
 CC and the pbp1A gene. The products and methods can be used for detecting
 CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
 CC used for simultaneously diagnosing pneumococcal meningitis and
 CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.

